RECEPTOR DEORPHANIZATION USING TAGGED MOLECULAR LIBRARIES

This application claims priority from U.S. provisional application Ser. No. 60/440,840 filed
5   17 January 2003, U.S. provisional application Ser. No. 60/440,838 filed 17 January 2003, and U.S.
application Ser. No. 10/379,965 filed 4 March 2003, all of which are incorporated herein by
reference.

FIELD OF THE INVENTION

10   This invention relates to methods and compositions for identifying classes of compounds
that specifically bind to cell surface receptors.

BACKGROUND OF THE INVENTION

G-protein coupled receptors (GPCRs) represent one of the most important families of drug
15   targets. G protein-mediated signaling systems have been identified in many divergent organisms,
such as mammals and yeast. GPCRs respond to, among other extracellular signals,
neurotransmitters, hormones, odorants and light. GPCRs are thought to represent a large
superfamily of proteins that are characterized by the seven distinct hydrophobic regions, each about
20-30 amino acids in length, that forms the transmembrane domain. The amino acid sequence is
20   not conserved across the entire superfamily, but each phylogenetically related subfamily contains a
number of highly conserved amino acid motifs that can be used to identify and classify new
members. Individual GPCRs activate particular signal transduction pathways, although at least ten
different signal transduction pathways are known to be activated via GPCRs. For example, the
beta 2-adrenergic receptor (βAR) is a prototype mammalian GPCR. In response to agonist binding,
25   βAR receptors activate a G protein (G_s) which in turn stimulates adenylate cyclase and cyclic
adenosine monophosphate production in the cell.

It has been postulated that members of the GPCR superfamily desensitize via a common
mechanism involving G protein-coupled receptor kinase (GRK) phosphorylation followed by
arrestin binding. The protein β-arrestin regulates GPCR signal transduction by binding agonist-
30   activated receptors that have been phosphorylated by G protein receptor kinases. The β-arrestin
protein remains bound to the GPCR during receptor internalization. The interaction between a
GPCR and α-arrestin can be measured using several methods. In one example, the β-arrestin
protein is fused to green fluorescent protein to create a protein fusion (Barak et al. (1997) *J. Biol.
Chem.* 272(44):27497-500). The agonist-dependent binding of β-arrestin to a GPCR can be
35   visualized by fluorescence microscopy. Microscopy can also be used to visualize the subsequent
trafficking of the GPCR β-arrestin complex to clathrin coated pits. Other methods for measuring
binding of β-arrestin to a GPCR in live cells include techniques such as FRET (fluorescence

resonance energy transfer), BRET (bioluminescent energy transfer) or enzyme complementation (Rossi et al. (1997) *Proc. Natl Acad. Sci. U S A* 94(16):8405-10).

At present, there are nearly 400 GPCRs whose natural ligands and function are known. These known GPCRs, named for their endogenous ligands, have been classified into five major

5 categories: Class-A Rhodopsin-like; Class-B Secretin-like; Class-C Metabotropic glutamate/pheromone; Class-D Fungal pheromone; Class-E cAMP (dictyostelium). Representative members of Class-A are the amine receptors (e.g., muscarinic, nicotinic, adrenergic, adenosine, dopamine, histamine and serotonin), the peptide receptors (e.g., angiotensin, bradykinin, chemokines, endothelin and opioid), the hormone receptors (e.g., follicle stimulating, lutropin and

10 thyrotropin), and the sensory receptors, including rhodopsin (light), olfactory (smell) and gustatory (taste) receptors. Representatives of Class-B include secretin, calcitonin, gastrin and glucagon receptors.

Many available therapeutic drugs in use today target GPCRs, as they mediate vital physiological responses, including vasodilation, heart rate, bronchodilation, endocrine secretion,

15 and gut peristalsis (Wilson and Bergsma (2000) Pharm. News 7: 105-114). For example, ligands to β-adrenergic receptors are used in the treatment of anaphylaxis, shock, hypertension, hypotension, asthma and other conditions. Additionally, diseases can be caused by the occurrence of spontaneous activation of GPCRs, where a GPCR cellular response is generated in the absence of a ligand. Drugs that are antagonists of GPCRs decrease this spontaneous activity (a process

20 known as inverse agonism) are important therapeutic agents, e.g. George et al, Nature Reviews Drug Discovery, 1: 808-820 (2002); Howard et al, Trends in Pharmaceutical Sciences, 22: 132-140 (2001); Seymour, Current Drug Targets, 2: 117-133 (2001).

Summary of the Invention

25 Due to the therapeutic importance of GPCRs, methods for the rapid screening of compounds for GPCR ligand activity are desirable. The present invention provides a method of screening test compounds and test conditions for the ability to modulate (activate or inhibit, enhance or depress) a GPCR pathway, and provides methods of assessing GPCR pathway function, such as the function of an orphan GPCR, in a cell in general. In another aspect of the present

30 method, lipophilic photosensitizers are attached to the cellular membranes. A candidate ligand or a library of candidate ligands can be attached to a molecular, after which the ligand is allowed to bind to the receptor. After excitation of the photosensitizer with a light source, the cleavable linker is cleaved releasing the molecular tag. The released molecular tag can be detected in the extracellular fluid, as detailed above, which provides information on the structure of the ligand for

35 the GPCR.

In one aspect, the present invention provides methods for screening modulators of GPCR activity comprising: a) providing a cell expressing a known or unknown GPCR, wherein the cell is

2

labeled with a lipophilic photosensitizer, b) exposing the cell to a test compound conjugated by a cleavable linkage to a molecular tag; c) illuminating the photosensitizer to generate singlet oxygen that cleaves the molecular tag, and d) detecting the signal from the released molecular tag. In another aspect, the method of the invention may further provide a step of comparing the signal

5   produced in the presence of the test compound with the signal produced in the absence, wherein changes in the signal indicates that the compound is a modulator of a GPCR.

The present invention thus provides a convenient method of identifying modulators for an orphan GPCR. Orphan GPCRs are novel receptors typically identified by sequence comparison-based methods, but whose cognate ligands are not known. It is estimated that from 400 to as many as

10   5000 orphan GPCRs may be coded for in the human genome, representing a vast potential for developing new drugs.

Brief Description of the Drawings

Figures 1A-1B illustrates one aspect of the method of the invention wherein a peptide

15   library having a molecular tag is applied to cells expression an orphan GPCR.

Figure 1C illustrates one embodiment of a peptide library for determining structural information about an orphan GPCR through peptide binding measurements.

Figure 1D illustrate a method of anchoring photosensitizers to the cell surface membrane of a cell.

20   Figure 2 illustrates one exemplary synthetic approach starting with commercially available 6-carboxy fluorescein, where the phenolic hydroxyl groups are protected using an anhydride. Upon standard extractive workup, a 95% yield of product is obtained. This material is phosphitylated to generate the phosphoramidite monomer.

Figure 3 illustrates the use of a symmetrical bis-amino alcohol linker as the amino alcohol

25   with the second amine then coupled with a multitude of carboxylic acid derivatives.

Figure 4 shows the structure of several benzoic acid derivatives that can serve as mobility modifiers.

Figure 5 illustrates the use of an alternative strategy that uses 5-aminofluorescein as starting material and the same series of steps to convert it to its protected phosphoramidite

30   monomer.

Figure 6 illustrates several amino alcohols and diacid dichlorides that can be assembled into mobility modifiers in the synthesis of molecular tags.

Figures 7 A-F illustrate oxidation-labile linkages and their respective cleavage reactions mediated by singlet oxygen.

Figures 8 A-B illustrate the general methodology for conjugation of an e-tag moiety to an antibody to form an e-tag probe, and the reaction of the resulting probe with singlet oxygen to produce a sulfinic acid moiety as the released molecular tag.

Figures 9A-J show the structures of e-tag moieties that have been designed and synthesized. (Pro1 is commercially available from Molecular Probes, Inc.)

Figures 10 A-I illustrate the chemistries of synthesis of the e-tag moieties illustrated in Figure 9.

Definitions

"Membrane-associated analyte" means a substance, compound, molecule, or component or part of any of the foregoing that has, or is attached to another molecule that has, a lipophilic moiety capable of anchoring it in a membrane. A membrane-associeated analyte is in a sample and include but are not limited to peptides, proteins, polynucleotides, polypeptides, oligonucleotides, organic molecules, haptens, epitopes, parts of biological cells, posttranslational modifications of proteins, receptors, complex sugars, vitamins, hormones, and the like. There may be more than one analyte associated with a single molecular entity, e.g. different phosphorylation sites on the same protein.

"Antibody" means an immunoglobulin that specifically binds to, and is thereby defined as complementary with, a particular spatial and polar organization of another molecule. The antibody can be monoclonal or polyclonal and can be prepared by techniques that are well known in the art such as immunization of a host and collection of sera (polyclonal) or by preparing continuous hybrid cell lines and collecting the secreted protein (monoclonal), or by cloning and expressing nucleotide sequences or mutagenized versions thereof coding at least for the amino acid sequences required for specific binding of natural antibodies. Antibodies may include a complete immunoglobulin or fragment thereof, which immunoglobulins include the various classes and isotypes, such as IgA, IgD, IgE, IgG1, IgG2a, IgG2b and IgG3, IgM, etc. Fragments thereof may include Fab, Fv and F(ab')2, Fab', and the like. In addition, aggregates, polymers, and conjugates of immunoglobulins or their fragments can be used where appropriate so long as binding affinity for a particular polypeptide is maintained.

"Antibody binding composition" means a molecule or a complex of molecules that comprise one or more antibodies and derives its binding specificity from an antibody. Antibody binding compositions include, but are not limited to, antibody pairs in which a first antibody binds specifically to a target molecule and a second antibody binds specifically to a constant region of the first antibody; a biotinylated antibody that binds specifically to a target molecule and streptavidin derivatized with moieties such as molecular tags or photosensitizers; antibodies specific for a target molecule and conjugated to a polymer, such as dextran, which, in turn, is derivatized with moieties such as molecular tags or photosensitizers; antibodies specific for a target molecule and conjugated

4

to a bead, or microbead, or other solid phase support, which, in turn, is derivatized with moieties such as molecular tags or photosensitizers, or polymers containing the latter.

"Binding compound" means any molecule to which molecular tags can be directly or indirectly attached that is capable of specifically binding to a membrane-associated analyte.

5 Binding compounds include, but are not limited to, antibodies, antibody binding compositions, peptides, proteins, particularly secreted proteins and orphan secreted proteins, nucleic acids, and organic molecules having a molecular weight of up to 1000 daltons and consisting of atoms selected from the group consisting of hydrogen, carbon, oxygen, nitrogen, sulfur, and phosphorus.

"Capillary-sized" in reference to a separation column means a capillary tube or channel in a

10 plate or microfluidics device, where the diameter or largest dimension of the separation column is between about 25-500 microns, allowing efficient heat dissipation throughout the separation medium, with consequently low thermal convection within the medium.

"Chromatography" or "chromatographic separation" as used herein means or refers to a method of analysis in which the flow of a mobile phase, usually a liquid, containing a mixture of

15 compounds, e.g. including analytes, promotes the separation of such compounds by a differential distribution between the mobile phase and a stationary phase, usually a solid.

A "separation profile" in reference to the separation of molecular tags means a chart, graph, curve, bar graph, or other representation of signal intensity data versus time, or other variable related to time, that provides a readout, or measure, of the number of molecular tags of

20 each type produced in an assay. A separation profile may be an electropherogram, a chromatogram, an electrochromatogram, or like graphical representations of data depending on the separation technique employed. A "peak" or a "band" or a "zone" in reference to a separation profile means a region where a separated compound is concentrated. There may be multiple separation profiles for a single assay if, for example, different molecular tags have different

25 fluorescent labels having distinct emission spectra and data is collected and recorded at multiple wavelengths.

"Specific" or "specificity" in reference to the binding of one molecule to another molecule, such as a probe for a target polynucleotide, means the recognition, contact, and formation of a stable complex between the two molecules, together with substantially less recognition, contact, or

30 complex formation of that molecule with other molecules. In one aspect, "specific" in reference to the binding of a first molecule to a second molecule means that to the extent the first molecule recognizes and forms a complex with another molecules in a reaction or sample, it forms the largest number of the complexes with the second molecule. Preferably, this largest number is at least fifty percent. Generally, molecules involved in a specific binding event have areas on their surfaces or

35 in cavities giving rise to specific recognition between the molecules binding to each other. Examples of specific binding include antibody-antigen interactions, enzyme-substrate interactions, formation of duplexes or triplexes among polynucleotides and/or oligonucleotides, receptor-ligand

5

interactions, and the like. As used herein, "contact" in reference to specificity or specific binding means two molecules are close enough that weak noncovalent chemical interactions, such as Van der Waal forces, hydrogen bonding, ionic and hydrophobic interactions, and the like, dominate the interaction of the molecules. As used herein, "stable complex" in reference to two or more

5 molecules means that such molecules form noncovalently linked aggregates, e.g. by specific binding, that under assay conditions are thermodynamically more favorable than a non-aggregated state.

As used herein, the term "spectrally resolvable" in reference to a plurality of fluorescent labels means that the fluorescent emission bands of the labels are sufficiently distinct, *i.e.*

10 sufficiently non-overlapping, that molecular tags to which the respective labels are attached can be distinguished on the basis of the fluorescent signal generated by the respective labels by standard photodetection systems, *e.g.* employing a system of band pass filters and photomultiplier tubes, or the like, as exemplified by the systems described in U.S. Pat. Nos. 4,230,558; 4,811,218, or the like, or in Wheeless et al, pgs. 21-76, in Flow Cytometry: Instrumentation and Data Analysis

15 (Academic Press, New York, 1985).

"Normal phase" in reference to chromatographic separation means that separation operates on the basis of hydrophilicity and lipophilicity by using a polar stationary phase and a less polar mobile phase. Thus hydrophobic compounds elute more quickly than do hydrophilic compounds. Exemplary groups on a solid phase for normal phase chromatography are amine ($-NH_2$) and

20 hydroxyl (-OH) groups.

"Reverse phase" in reference to chromatographic separation means that separation operates on the basis of hydrophilicity and lipophilicity. The stationary phase usually consists of silica based packings with n-alkyl chains or phenyl groups covalently bound. For example, C-8 signifies an octyl chain and C-18 an octadecyl ligand in the matrix. The more hydrophobic the matrix on each

25 ligand, the greater is the tendancy of the column to retain hydrophobic moieties. Thus hydrophilic compounds elute more quickly than do hydrophobic compounds.

"Ion-exchange" in reference to chromatographic separation means that separation operates on the basis of selective exchange of ions in the sample with counterions in the stationary phase. Ion exchange is performed with columns containing charge-bearing functional groups attached to a

30 polymer matrix. The functional ions are permanently bonded to the column and each has a counterion attached. The sample is retained by replacing the counterions of the stationary phase with its own ions. The sample is eluted from the column by changing the properties of the mobile phase do that the mobile phase will now displace the sample ions from the stationary phase, (ie. changing the pH).

The term "sample" in the present specification and claims is used in a broad sense. On the one hand it is meant to include a specimen or culture (e.g., microbiological cultures). On the other hand, it is meant to include both biological and environmental samples. A sample may include a specimen of synthetic origin. Biological samples may be animal, including human, fluid, solid
5   (e.g., stool) or tissue, as well as liquid and solid food and feed products and ingredients such as dairy items, vegetables, meat and meat by-products, and waste. Biological samples may include materials taken from a patient including, but not limited to cultures, blood, saliva, cerebral spinal fluid, pleural fluid, milk, lymph, sputum, semen, needle aspirates, and the like. Biological samples may be obtained from all of the various families of domestic animals, as well as feral or wild
10   animals, including, but not limited to, such animals as ungulates, bear, fish, rodents, etc. Environmental samples include environmental material such as surface matter, soil, water and industrial samples, as well as samples obtained from food and dairy processing instruments, apparatus, equipment, utensils, disposable and non-disposable items. These examples are not to be construed as limiting the sample types applicable to the present invention.
15   As used herein, the term "kit" refers to any delivery system for delivering materials. In the context of reaction assays, such delivery systems include systems that allow for the storage, transport, or delivery of reaction reagents (e.g., probes, enzymes, etc. in the appropriate containers) and/or supporting materials (e.g., buffers, written instructions for performing the assay etc.) from one location to another. For example, kits include one or more enclosures (e.g., boxes) containing
20   the relevant reaction reagents and/or supporting materials. Such contents may be delivered to the intended recipient together or separately. For example, a first container may contain an enzyme for use in an assay, while a second container contains probes.

"Polypeptide" refers to a class of compounds composed of amino acid residues chemically bonded together by amide linkages with elimination of water between the carboxy group of one
25   amino acid and the amino group of another amino acid. A polypeptide is a polymer of amino acid residues, which may contain a large number of such residues. Peptides are similar to polypeptides, except that, generally, they are comprised of a lesser number of amino acids. Peptides are sometimes referred to as oligopeptides. There is no clear-cut distinction between polypeptides and peptides. For convenience, in this disclosure and claims, the term "polypeptide" will be used to
30   refer generally to peptides and polypeptides. The amino acid residues may be natural or synthetic.

"Protein" refers to a polypeptide, usually synthesized by a biological cell, folded into a defined three-dimensional structure. Proteins are generally from about 5,000 to about 5,000,000 or more in molecular weight, more usually from about 5,000 to about 1,000,000 molecular weight, and may include posttranslational modifications, such acetylation, acylation, ADP-ribosylation,
35   amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond

7

formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, phosphorylation, prenylation, racemization, selenoylation, sulfation, and ubiquitination, e.g. Wold, F., Post-translational Protein

5   Modifications: Perspectives and Prospects, pgs. 1-12 in Post-translational Covalent Modification of Proteins, B. C. Johnson, Ed., Academic Press, New York, 1983. Proteins include, by way of illustration and not limitation, cytokines or interleukins, enzymes such as, *e.g.*, kinases, proteases, galactosidases and so forth, protamines, histones, albumins, immunoglobulins, scleroproteins, phosphoproteins, mucoproteins, chromoproteins, lipoproteins, nucleoproteins, glycoproteins, T-cell

10   receptors, proteoglycans, unclassified proteins, *e.g.*, somatotropin, prolactin, insulin, pepsin, proteins found in human plasma, blood clotting factors, blood typing factors, protein hormones, cancer antigens, tissue specific antigens, peptide hormones, nutritional markers, tissue specific antigens, and synthetic peptides.

The term "secreted protein" refers to proteins that are (i) expressed intracellularly, (ii)

15   secreted from the cell into the extracellular medium, *e.g.*, typically requiring a leader sequence that directs the expressed protein from the endoplasmic reticulum through the cell membrane, and (iii) act on a receptor, typically a cell-surface receptor, to effect or initiate some cellular event or activity, which may be an intracellular event, including cell proliferation or stimulation, a cell-surface event, or cell-cell interaction event.

20   The term "ligand" is also used herein to refer to a secreted protein or protein thereof which binds to a given receptor, through a ligand-receptor interaction.

The term "orphan secreted protein" means a secreted protein (or ligand) which is uncharacterized as to one or more of (i) the cell type to which the protein binds, (ii) the receptor to which the protein binds, and (iii) the action produced by the binding of the protein to its receptor.

25   Examples include cytokines and lymphokines, including those capable of stimulating production and/or proliferation of spleen cells, lymph node cells or thymocytes, proteins that exhibit immune stimulating or immune suppressing activity, proteins regulating hematopoisesis, tissue growth, cell chemotactic or chemokinetic events, such as cell adhesion molecules, or cell-recruitment ligands, proteins with anti-Inflammatory activity and anti-tumor activity.

30

DETAILED DESCRIPTION OF THE INVENTION

The invention provides a method for structurally identifying a ligand for a cell surface receptor having a known or unknown function. In one aspect, biological cells expressing a receptor of interest are treated with a lipophilic sensitizer to form sensitizer-treated membranes on their

35   surfaces. A plurality of libraries of binding compounds each with a different molecular tag are then incubated with the cells under binding conditions, so that members of the libraries having higher affinities for the receptor of interest bind to the receptor in higher quantities. Optionally, unbound,

or weakly bound, binding compound is removed by washing, after which the sensitizers of the sensitizer-treated membranes are induced to generate an active species for cleaving the cleavable linkage attaching molecular tags to their respective binding compounds. The released molecular tags are then separated and identified to provide a measure of the binding strength of members of

5   the respective libraries. In one aspect such binding compounds are peptides, preferably having a length in the range of from six to fifteen amino acids, and more preferably, having a length in the range of eight to twelve amino acids.

One embodiment of the invention is illustrated in Figures 1A and 1B. Cells (100) having cell surface membrane (101) and receptors (102, "R1"), (104, "R2"), and (106, "R3")—the

10 membrane-associated analytes--are incubated with lipophilic photosensitizers (108) to form (107) photosensitizer-treated membranes (109), which are part of intact cells (110). Binding compounds (112) having molecular tags ("mT_k") attached are combined with photosensitizer-treated membranes (109) under conditions that permit specific binding of binding compounds (112) to their respective target membrane-associated analytes. Reaction mixture (116) is then illuminated

15 (118) with a light of wavelength and intensity to excite the membrane-bound photosensitizers to generated singlet oxygen which, in turn, cleaves the molecular tags (122) from the binding compounds, shown in Fig. 1B. Molecular tags (122) are then separated (124) from reaction mixture (120) and identified in separation profile (126). The relative amounts of bound peptide, as measured by the molecular tag peaks, provides information on the nature of receptors on the cell

20 surfaces and to the identification of peptides that bind to the receptors.

In one aspect, tagged peptide libraries have the form illustrated in Fig. 1C. A peptide having a predetermined length, N, and amino acid sequence, AA_1-AA_2-AA_3- ... -AA_N, is selected. Next, a set of K peptide mixtures is synthesized each having the same molecular tag, mT_k. The number K may vary widely and be larger than or smaller than N. In one aspect, K is equal to N and

25 at each amino acid position a mixture is created by inserting an amino acid randomly from a defined set of amino acids, as illustrated by the "XXi"(130, 132, and 134) in Fig. 1C. Usually, the defined set is the natural amino acids, e.g. having a size of about 20, but it may also include unnatural amino acids and be larger, e.g. having a size of 30 or more. The defined set may also be a subset of the natural amino acids, e.g. those of some common physiochemical property, such as

30 charged amino acids, lipophilic amino acids, or the like. As mentioned above, in one aspect, N is in the range of from 8 to 12. Molecular tags may be conveniently attached via the N-terminus of the peptide using conventional chemistries, e.g. either on-column or off-column attachment via an NHS-ester of the molecular tags, as disclosed below.

As described more fully below, target membrane-associated analytes are determined by

35 separation and identification of the released molecular tags. A wide variey of separation techniques may be employed that can distinguish molecules based on one or more physical, chemical, or optical differences amoung molecules being separated including but not limited to

9

electrophoretic mobility, molecular weight, shape, solubility, pKa, hydrophobicity, charge, charge/mass ratio, polarity, or the like. In one aspect, molecular tags in a plurality differ in electrophoretic mobility and optical detection characteristics and are separated by electrophoresis. In another aspect, molecular tags in a plurality differ in molecular weight, shape, solubility, pKa,

5   hydrophobicity, charge, polarity, and are separated by normal phase or reverse phase HPLC, ion exchange HPLC, capillary electrochromatography, or like technique.

Another aspect of the present invention is providing sets of molecular tags that may be separated into distinct bands or peaks by the separation technique employed after they are released from microparticles. Molecular tags within a set may be chemically diverse; however, for

10 convenience, sets of molecular tags are usually chemically related. For example, they may all be peptides, or they may consist of different combinations of the same basic building blocks or monomers, or they may be synthesized using the same basic scaffold with different substituent groups for imparting different separation characteristics, as described more fully below. The number of molecular tags in a plurality may vary depending on several factors including the mode

15 of separation employed, the labels used on the molecular tags for detection, the sensitivity of the binding moieties, the efficiency with which the cleavable linkages are cleaved, and the like. In one aspect, the number of molecular tags in a plurality ranges from 2 to several tens, e.g. 50. In other aspects, the size of the plurality may be in the range of from 2 to 40, 2 to 20, 2 to 10, 3 to 50, 3 to 20, 3 to 10, 4 to 50, 4 to 10, 5 to 20, or 5 to 10.

20

Molecular Tags and Cleavable Linkages

In one embodiment, molecular tags are cleaved from a binding compound by reaction of a cleavable linkage with an active species, such as singlet oxygen, generated by a cleavage-inducing moiety, e.g. Singh et al, International patent publication WO 01/83502.

25 An aspect of the invention includes providing a plurality of libraries binding compounds, wherein each different library has one or more molecular tags attached through cleavable linkages. The nature of the libraries of binding compounds, cleavable linkage and molecular tag may vary widely. In one aspect, libraries for use with the invention are disjoint subsets of a larger library of compounds. Preferably, compounds in such libraries are linear oligomers of N monomers capable

30 of combinatorial systhesis. Disjoint subsets may be conveniently synthesized by selecting a starting sequence of monomers following by generation of libraries by successively incorporating a random selection of monomers at each position along the sequence of the starting sequence. Thus, for example, for a 10-mer starting sequence and an 8 monomer selection, ten 10-mer libraries could be produced each having 8 members. Preferably, the libraries of binding compounds are peptide

35 libraries, as illustrated in Fig. 1A and 1B.

In one aspect, a molecular tag attached to a binding compound can be represented by the formula:

10

wherein P is a peptide library; L is a cleavable linkage; and E is a molecular tag. Preferably,

5   cleavable linkage, L, is an oxidation-labile linkage, and more preferably, it is a linkage that may be cleaved by singlet oxygen. Cleavable linkages, e.g. oxidation-labile linkages, and molecular tags, E, are attached to P by way of conventional chemistries.

      When L is oxidation labile, L is preferably a thioether or its selenium analog; or an olefin, which contains carbon-carbon double bonds, wherein cleavage of a double bond to an oxo group,

10   releases the molecular tag, E. Illustrative olefins include vinyl sulfides, vinyl ethers, enamines, imines substituted at the carbon atoms with an α-methine (CH, a carbon atom having at least one hydrogen atom), where the vinyl group may be in a ring, the heteroatom may be in a ring, or substituted on the cyclic olefinic carbon atom, and there will be at least one and up to four heteroatoms bonded to the olefinic carbon atoms. The resulting dioxetane may decompose

15   spontaneously, by heating above ambient temperature, usually below about 75°C, by reaction with acid or base, or by photo-activation in the absence or presence of a photosensitizer. Such reactions are described in the following exemplary references: Adam and Liu, J. Amer. Chem. Soc. 94, 1206-1209, 1972, Ando, *et al.*, J.C.S. Chem. Comm. 1972, 477-8, Ando, *et al.*, Tetrahedron 29, 1507-13, 1973, Ando, *et al.*, J. Amer. Chem. Soc. 96, 6766-8, 1974, Ando and Migita, ibid. 97,

20   5028-9, 1975, Wasserman and Terao, Tetra. Lett. 21, 1735-38, 1975, Ando and Watanabe, ibid. 47, 4127-30, 1975, Zaklika, *et al.*, Photochemistry and Photobiology 30, 35-44, 1979, and Adam, *et al.*, Tetra. Lett. 36, 7853-4, 1995. See also, U.S. Patent no. 5,756,726.

      The formation of dioxetanes is obtained by the reaction of singlet oxygen with an activated olefin substituted with an molecular tag at one carbon atom and the binding moiety at the other

25   carbon atom of the olefin. See, for example, U.S. Patent No. 5,807,675. These cleavable linkages may be depicted by the following formula:

$$-W-(X)_nC_\alpha= C_\beta(Y)(Z)-$$
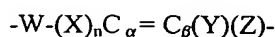
wherein:

30   W may be a bond, a heteroatom, *e.g.*, O, S, N, P, M (intending a metal that forms a stable covalent bond), or a functionality, such as carbonyl, imino, etc., and may be bonded to X or C_α; at least one X will be aliphatic, aromatic, alicyclic or heterocyclic and bonded to C_α through a hetero atom, *e.g.*, N, O, or S and the other X may be the same or different and may in addition be hydrogen, aliphatic, aromatic, alicyclic or heterocyclic, usually being aromatic or aromatic

35   heterocyclic wherein one X may be taken together with Y to form a ring, usually a heterocyclic ring, with the carbon atoms to which they are attached, generally when other than hydrogen being from about 1 to 20, usually 1 to 12, more usually 1 to 8 carbon atoms and one X will have 0 to 6,

usually 0 to 4 heteroatoms, while the other X will have at least one heteroatom and up to 6

heteroatoms. usually 1 to 4 heteroatoms;

Y will come within the definition of X, usually being bonded to C_β through a heteroatom

and as indicated may be taken together with X to form a heterocyclic ring;

5   Z will usually be aromatic, including heterocyclic aromatic, of from about 4 to 12, usually

4 to 10 carbon atoms and 0 to 4 heteroatoms, as described above, being bonded directly to C_β or

through a heteroatom, as described above;

n is 1 or 2, depending upon whether the molecular tag is bonded to C_α or X;

wherein one of Y and Z will have a functionality for binding to the binding moiety, or be

10   bound to the binding moiety, e.g. by serving as, or including a linkage group, to a binding moiety,

T.

Preferably, W, X, Y, and Z are selected so that upon cleavage molecular tag, E, is within

the size limits described below.

Illustrative cleavable linkages include S(molecular tag)-3-thiolacrylic acid, N(molecular

15   tag), N-methyl 4-amino-4-butenoic acid, 3-hydroxyacrolein, N-(4-carboxyphenyl)-2-(molecular

tag)-imidazole, oxazole, and thiazole.

Also of interest are N-alkyl acridinyl derivatives, substituted at the 9 position with a

divalent group of the formula:

$$- (CO) X^1 (A) -$$
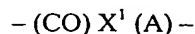
20   wherein:

X^1 is a heteroatom selected from the group consisting of O, S, N, and Se, usually one of the

first three; and

A is a chain of at least 2 carbon atoms and usually not more than 6 carbon atoms

substituted with an molecular tag, where preferably the other valences of A are satisfied by

25   hydrogen, although the chain may be substituted with other groups, such as alkyl, aryl, heterocyclic

groups, etc., A generally being not more than 10 carbon atoms.

Also of interest are heterocyclic compounds, such as diheterocyclopentadienes, as

exemplified by substituted imidazoles, thiazoles, oxazoles, etc., where the rings will usually be

substituted with at least one aromatic group and in some instances hydrolysis will be necessary to

30   release the molecular tag.

Also of interest are tellurium (Te) derivatives, where the Te is bonded to an ethylene group

having a hydrogen atom β to the Te atom, wherein the ethylene group is part of an alicyclic or

heterocyclic ring, that may have an oxo group, preferably fused to an aromatic ring and the other

valence of the Te is bonded to the molecular tag. The rings may be coumarin, benzoxazine,

35   tetralin, etc.

Several preferred cleavable linkages and their cleavage products are illustrated in Figures 7

A-F. The thiazole cleavable linkage, "-CH_2-thiazole-$(CH2)_n$-C(=O)-NH-protein," shown in Fig.

12

7A, results in an molecular tag with the moiety "-CH$_2$-C(=O)-NH-CHO." Preferably, n is in the range of from 1 to 12, and more preferably, from 1 to 6. The oxazole cleavable linkage, "-CH$_2$-oxazole-(CH2)$_n$-C(=O)-NH-protein," shown in Fig. 7B, results in an molecular tag with the moiety "-CH$_2$-C(=O)O-CHO." An olefin cleavable linkage (Fig. 7C) is shown in connection with the

5   binding compound embodiment "P-L-M-D," described above and with D being a fluorescein dye. The olefin cleavable linkage may be employed in other embodiments also. Cleavage of the illustrated olefin linkage results in an molecular tag of the form: "R-(C=O)-M-D," where "R" may be any substituent within the general description of the molecular tags, E, provided above. Preferably, R is an electron-donating group, e.g. Ullman et al, U.S. patent 6,251,581; Smith and

10   March, March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure, 5th Edition (Wiley-Interscience, New York, 2001); and the like. More preferably, R is an electron-donating group having from 1-8 carbon atoms and from 0 to 4 heteroatoms selected from the group consisting of O, S, and N. In further preference, R is -N(Q)$_2$, -OQ, p-[C$_6$H$_4$N(Q)$_2$], furanyl, n-alkylpyrrolyl, 2-indolyl, or the like, where Q is alkyl or aryl. In further reference to the olefin

15   cleavable linkage of Fig. 7C, substituents "X" and "R" are equivalent to substituents "X" and "Y" of the above formula describing cleavable linkage, L. In particular, X in Fig. 7C is preferably morpholino, -OR', or -SR", where R' and R" are aliphatic, aromatic, alicyclic or heterocyclic having from 1 to 8 carbon atoms and 0 to 4 heteroatoms selected from the group consisting of O, S. and N. A preferred thioether cleavable linkage is illustrated in Fig. 6D having the form "-(CH$_2$)$_2$-S-

20   CH(C$_6$H$_5$)C(=O)NH-(CH$_2$)$_n$-NH-," wherein n is in the range of from 2 to 12, and more preferably, in the range of from 2 to 6. Thioether cleavable linkages of the type shown in Fig. 7D may be attache to binding moieties, T, and molecular tags, E, by way of precursor compounds shown in Figures 7E and 7F. To attach to an amino group of a binding moiety, T, the terminal hydroxyl is converted to an NHS ester by conventional chemistry. After reaction with the amino group and

25   attachment, the Fmoc protection group is removed to produce a free amine which is then reacted with an NHS ester of the molecular tag, such as compounds produced by the schemes of Figures 1, 2, and 4, with the exception that thelast reaction step is the addition of an NHS ester, instead of a phosphoramidite group.

      Molecular tag, E, is a water soluble organic compound that is stable with respect to the

30   active species, especially singlet oxygen, and that includes a detection or reporter group. Otherwise, E may vary widely in size and structure. In one aspect, E has a molecular weight in the range of from about 100 to about 2500 daltons, more preferably, from about 100 to about 1500 daltons. Preferred structures of E are described more fully below. The detection group may generate an electrochemical, fluorescent, or chromogenic signal. Preferably, the detection group

35   generates a fluorescent signal.

      Molecular tags within a plurality of a composition each have either a unique separation characteristic and/or a unique optical property with respect to the other members of the same

<div align="center">13</div>

plurality. In one aspect, the chromatographic or electrophoretic separation characteristic is retention time under set of standard separation conditions conventional in the art, e.g. voltage, column pressure, column type, mobile phase, electrophoretic separation medium, or the like. In another aspect, the optical property is a fluorescence property, such as emission spectrum,

5   fluorescence lifetime, fluorescence intensity at a given wavelength or band of wavelengths, or the like. Preferably, the fluorescence property is fluorescence intensity. For example, each molecular tag of a plurality may have the same fluorescent emission properties, but each will differ from one another by virtue of a unique retention time. On the other hand, or two or more of the molecular tags of a plurality may have identical retention times, but they will have unique fluorescent

10   properties, e.g. spectrally resolvable emission spectra, so that all the members of the plurality are distinguishable by the combination of molecular separation and fluorescence measurement.

In one aspect, molecular tag, E, is (M, D), where M is a mobility-modifying moiety and D is a detection moiety. The notation "(M, D)" is used to indicate that the ordering of the M and D moieties may be such that either moiety can be adjacent to the cleavable linkage, L. That is, "P-L-

15   (M, D)" designates binding compound of either of two forms: "P-L-M-D" or "P-L-D-M."

Detection moiety, D, may be a fluorescent label or dye, a chromogenic label or dye, an electrochemical label, or the like. Preferably, D is a fluorescent dye. Exemplary fluorescent dyes for use with the invention include water-soluble rhodamine dyes, fluoresceins, 4,7-dichlorofluoresceins, benzoxanthene dyes, and energy transfer dyes, disclosed in the following references: Handbook of

20   Molecular Probes and Research Reagents, 8th ed., (Molecular Probes, Eugene, 2002); Lee et al, U.S. patent 6,191,278; Lee et al, U.S. patent 6,372,907; Menchen et al, U.S. patent 6,096,723; Lee et al, U.S. patent 5,945,526; Lee et al, Nucleic Acids Research, 25: 2816-2822 (1997); Hobb, Jr., U.S. patent 4,997,928; Khanna et al., U.S. patent 4,318,846; Reynolds, U.S. patent 3,932,415; Eckert et al, U.S. patent 2,153,059; Eckert et al, U.S. patent 2,242,572; Taing et al, International patent publication

25   WO 02/30944; and the like. Further specific exemplary fluorescent dyes include 5- and 6-carboxyrhodamine 6G; 5- and 6-carboxy-X-rhodamine, 5- and 6-carboxytetramethylrhodamine, 5- and 6-carboxyfluorescein, 5- and 6-carboxy-4,7-dichlorofluorescein, 2',7'-dimethoxy-5- and 6-carboxy-4,7-dichlorofluorescein, 2',7'-dimethoxy-4',5'-dichloro-5- and 6-carboxyfluorescein, 2',7'-dimethoxy-4',5'-dichloro-5- and 6-carboxy-4,7-dichlorofluorescein, 1',2',7',8'-dibenzo-5- and 6-

30   carboxy-4,7-dichlorofluorescein, 1',2',7',8'-dibenzo-4',5'-dichloro-5- and 6-carboxy-4,7-dichlorofluorescein, 2',7'-dichloro-5- and 6-carboxy-4,7-dichlorofluorescein, and 2',4',5',7'-tetrachloro-5- and 6-carboxy-4,7-dichlorofluorescein. Most preferably, D is a fluorescein or a fluorescein derivative.

The size and composition of mobility-modifying moiety, M, can vary from a bond to about

35   100 atoms in a chain, usually not more than about 60 atoms, more usually not more than about 30 atoms, where the atoms are carbon, oxygen, nitrogen, phosphorous, boron and sulfur. Generally, when other than a bond, the mobility-modifying moiety has from about 0 to about 40, more usually

14

from about 0 to about 30 heteroatoms, which in addition to the heteroatoms indicated above may include halogen or other heteroatom. The total number of atoms other than hydrogen is generally fewer than about 200 atoms, usually fewer than about 100 atoms. Where acid groups are present, depending upon the pH of the medium in which the mobility-modifying moiety is present, various

5   cations may be associated with the acid group. The acids may be organic or inorganic, including carboxyl, thionocarboxyl, thiocarboxyl, hydroxamic, phosphate, phosphite, phosphonate, phosphinate, sulfonate, sulfinate, boronic, nitric, nitrous, etc. For positive charges, substituents include amino (includes ammonium), phosphonium, sulfonium, oxonium, etc., where substituents are generally aliphatic of from about 1 - 6 carbon atoms, the total number of carbon atoms per

10 heteroatom, usually be less than about 12, usually less than about 9. The side chains include amines, ammonium salts, hydroxyl groups, including phenolic groups, carboxyl groups, esters, amides, phosphates, heterocycles. M may be a homo-oligomer or a hetero-oligomer, having different monomers of the same or different chemical characteristics, e.g., nucleotides and amino acids.

15   In another aspect, (M,D) moieties are constructed from chemical scaffolds used in the generation of combinatorial libraries. For example, the following references describe scaffold compound useful in generating diverse mobility modifying moieties: peptoids (PCT Publication No WO 91/19735, Dec. 26, 1991), encoded peptides (PCT Publication WO 93/20242, Oct. 14 1993), random bio-oligomers (PCT Publication WO 92/00091, Jan. 9, 1992), benzodiazepines (U.S. Pat.

20 No. 5,288,514), diversomeres such as hydantoins, benzodiazepines and dipeptides (Hobbs DeWitt, S. et al., Proc. Nat. Acad. Sci. U.S.A. 90: 6909-6913 (1993), vinylogous polypeptides (Hagihara et al. J.Amer. Chem. Soc. 114: 6568 (1992)), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann, R. et al., J.Amer. Chem. Soc. 114: 9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen, C. et al. J.Amer. Chem. Soc. 116: 2661(1994)),

25 oligocarbamates (Cho, C. Y. et al. Science 261: 1303(1993)), peptidyl phosphonates (Campbell, D. A. et al., J. Org. Chem. 59:658(1994)); Cheng et al, U.S. patent 6,245,937; Heizmann et al, "Xanthines as a scaffold for molecular diversity," Mol. Divers. 2: 171-174 (1997); Pavia et al, Bioorg. Med. Chem., 4: 659-666 (1996); Ostresh et al, U.S. patent 5,856,107; Gordon, E. M. et al., J. Med. Chem. 37: 1385 (1994); and the like. Preferably, in this aspect, D is a substituent on a

30 scaffold and M is the rest of the scaffold.

In yet another aspect, (M, D) moieties are constructed from one or more of the same or different common or commercially available linking, cross-linking, and labeling reagents that permit facile assembly, especially using a commercial DNA or peptide synthesizer for all or part of the synthesis. In this aspect, (M, D) moieties are made up of subunits usually connected by

35 phosphodiester and amide bonds. Exemplary, precusors include, but are not limited to, dimethoxytrityl (DMT)-protected hexaethylene glycol phosphoramidite, 6-(4-Monomethoxytritylamino)hexyl-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite, 12-(4-

Monomethoxytritylamino)dodecyl-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite, 2-[2-(4-Monomethoxytrityl)aminoethoxy]ethyl-(2-cyanoethyl), N,N-diisopropyl)-phosphoramidite, (S-Trityl-6-mercaptohexyl)- (2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite, 5'-Fluorescein phosphoramidite, 5'-Hexachloro-Fluorescein Phosphoramidite, 5'-Tetrachloro-Fluorescein

5   Phosphoramidite, 9-O-Dimethoxytrityl-triethylene glycol,1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 3(4,4'Dimethoxytrityloxy)propyl-1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 5'-O-Dimethoxytrityl-1',2'-Dideoxyribose-3'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 18-O Dimethoxytritylhexaethyleneglycol,1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 12-(4,4'-Dimethoxytrityloxy)dodecyl-1-[(2-cyanoethyl)-(N,N-diisopropyl)]-

10   phosphoramidite, 1,3-bis-[5-(4,4'-dimethoxytrityloxy)pentylamido]propyl-2-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 1-[5-(4,4'-dimethoxytrityloxy)pentylamido]-3-[5-fluorenomethoxycarbonyloxy pentylamido]-propyl-2-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, Tris-2,2,2-[3-(4,4'-dimethoxytrityloxy)propyloxymethyl]ethyl-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite,

15   succinimidyl *trans*-4-(maleimidylmethyl) cyclohexane-1-carboxylate (SMCC), succinimidyl 3-(2-pyridyldithio)propionate (SPDP), succinimidyl acetylthioacetate, Texas Red-X-succinimidyl ester, 5- and 6-carboxytetramethylrhodamine succinimidyl ester, bis-(4-carboxypiperidinyl)sulfonerhodamine di(succinimidyl ester), 5- and 6-((N-(5-aminopentyl)aminocarbonyl)tetramethylrhodamine, succinimidyl 4-(p-maleimidophenyl)butyrate

20   (SMPB); N-γ-maleimidobutyryl-oxysuccinimide ester (GMBS); p-nitrophenyl iodoacetate (NPIA); 4-(4-N-maleimidophenyl)butyric acid hydrazide (MPBH); and like reagents. The above reagents are commercially available, e.g. from Glen Research (Sterling, VA), Molecular Probes (Eugene, OR), Pierce Chemical, and like reagent providers. Use of the above reagents in conventional synthetic schemes is well known in the art, e.g. Hermanson, Bioconjugate Techniques (Academic

25   Press, New York, 1996). In particular, M may be constructed from the following reagents: dimethoxytrityl (DMT)-protected hexaethylene glycol phosphoramidite, 6-(4-Monomethoxytritylamino)hexyl-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite, 12-(4-Monomethoxytritylamino)dodecyl-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite, 2-[2-(4-Monomethoxytrityl)aminoethoxy]ethyl-(2-cyanoethyl), N,N-diisopropyl)-phosphoramidite, (S-

30   Trityl-6-mercaptohexyl)- (2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite, 9-O-Dimethoxytrityl-triethylene glycol,1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 3(4,4'Dimethoxytrityloxy)propyl-1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 5'-O-Dimethoxytrityl-1',2'-Dideoxyribose-3'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 18-O Dimethoxytritylhexaethyleneglycol,1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 12-

35   (4,4'-Dimethoxytrityloxy)dodecyl-1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 1,3-bis-[5-(4,4'-dimethoxytrityloxy)pentylamido]propyl-2-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 1-[5-(4,4'-dimethoxytrityloxy)pentylamido]-3-[5-fluorenomethoxycarbonyloxy

16

pentylamido]-propyl-2-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, Tris-2,2,2-[3-(4,4'-dimethoxytrityloxy)propyloxymethyl]ethyl-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, succinimidyl *trans*-4-(maleimidylmethyl) cyclohexane-1-carboxylate (SMCC), succinimidyl 3-(2-pyridyldithio)propionate (SPDP), succinimidyl acetylthioacetate, succinimidyl 4-(p-

5   maleimidophenyl)butyrate (SMPB); N-γ-maleimidobutyryl-oxysuccinimide ester (GMBS); p-nitrophenyl iodoacetate (NPIA); and 4-(4-N-maleimidophenyl)butyric acid hydrazide (MPBH).

M may also comprise polymer chains prepared by known polymer subunit synthesis methods. Methods of forming selected-length polyethylene oxide-containing chains are well known, *e.g.* Grossman et al, U.S. patent 5,777,096. It can be appreciated that these methods, which

10   involve coupling of defined-size, multi-subunit polymer units to one another, directly or via linking groups, are applicable to a wide variety of polymers, such as polyethers (*e.g.*, polyethylene oxide and polypropylene oxide), polyesters (*e.g.*, polyglycolic acid, polylactic acid), polypeptides, oligosaccharides, polyurethanes, polyamides, polysulfonamides, polysulfoxides, polyphosphonates, and block copolymers thereof, including polymers composed of units of multiple subunits linked

15   by charged or uncharged linking groups. In addition to homopolymers, the polymer chains used in accordance with the invention include selected-length copolymers, *e.g.*, copolymers of polyethylene oxide units alternating with polypropylene units. As another example, polypeptides of selected lengths and amino acid composition (*i.e.*, containing naturally occurring or man-made amino acid residues), as homopolymers or mixed polymers.

20   In another aspect, after release, molecular tag, E, is defined by the formula:

A-M-D

wherein:

A is -C(=O)R, where R is aliphatic, aromatic, alicyclic or heterocyclic having from 1 to 8

25   carbon atoms and 0 to 4 heteroatoms selected from the group consisting of O, S. and N; $-CH_2-$ $C(=O)-NH-CHO$; $-SO_2H$; $-CH_2-C(=O)O-CHO$; $-C(=O)NH-(CH_2)_n-NH-C(=O)C(=O)-(C_6H_5)$, where n is in the range of from 2 to 12;

D is a fluorescent dye; and

M is as described above, with the proviso that the total molecular weight of A-M-D be

30   within the range of from about 100 to about 2500 daltons.

In another aspect, D is a fluorescein and the total molecular weight of A-M-D is in the range of from about 100 to about 1500 daltons.

In another aspect, M may be synthesized from smaller molecules that have functional groups that provide for linking of the molecules to one another, usually in a linear chain. Such

35   functional groups include carboxylic acids, amines, and hydroxy- or thiol- groups. In accordance with the present invention the charge-imparting moiety may have one or more side groups pending from the core chain. The side groups have a functionality to provide for linking to a label or to

17

another molecule of the charge-imparting moiety. Common functionalities resulting from the reaction of the functional groups employed are exemplified by forming a covalent bond between the molecules to be conjugated. Such functionalities are disulfide, amide, thioamide, dithiol, ether, urea, thiourea, guanidine, azo, thioether, carboxylate and esters and amides containing sulfur and

5   phosphorus such as, *e.g.*, sulfonate, phosphate esters, sulfonamides, thioesters, etc., and the like.

Attaching Molecular Tags to Binding Compounds

Extensive guidance can be found in the literature for covalently linking molecular tags to binding compounds, e.g. Hermanson, Bioconjugate Techniques, (Academic Press, New York,

10   1996), and the like. In one aspect of the invention, one or more molecular tags are attached directly or indirectly to common reactive groups on a binding compound. Common reactive groups include amine, thiol, carboxylate, hydroxyl, aldehyde, ketone, and the like, and may be coupled to molecular tags by commercially available cross-linking agents, e.g. Hermanson (cited above); Haugland, Handbook of Fluorescent Probes and Research Products, Ninth Edition

15   (Molecular Probes, Eugene, OR, 2002). In one embodiment, an NHS-ester of a molecular tag is reacted with a free amine on the binding compound.

Once each of the libraries of binding compounds is separately derivatized by a different molecular tag, it is pooled with other libraries to form a plurality of libraries of binding compounds. Usually, each different kind of library is present in a composition in the same

20   proportion; however, proportions may be varied as a design choice so that one or a subset of particular binding compounds are present in greater or lower proportion depending on the desirability or requirements for a particular embodiment or assay. Factors that may affect such design choices include, but are not limited to, affinity and avidity for a particular target, relative prevalence of a target, fluorescent characteristics of a detection moiety of a molecular tag, and

25   the like.

The chemistry for synthesizing peptides is well known in the art. See, for example, Marglin, *et al.*, Ann. Rev. Biochem. (1970) 39:841-866. In general, such syntheses involve blocking, with an appropriate protecting group, those functional groups that are not to be involved in the reaction. The free functional groups are then reacted to form the desired linkages. The

30   peptide can be produced on a resin as in the Merrifield synthesis (Merrifield, J. Am. Chem. Soc. (1980) 85:2149-2154 and Houghten *et al.*, Int. J. Pep. Prot. Res. (1980) 16:311-320. The peptide is then removed from the resin according to known techniques.

A summary of the many techniques available for the synthesis of peptides may be found in J. M. Stewart, *et al.*, "Solid Phase Peptide Synthesis, W. H. Freeman Co, San Francisco (1969); and

35   J. Meienhofer, "Hormonal Proteins and Peptides", (1973), vol. 2, p. 46, Academic Press (New York), for solid phase peptide synthesis; and E. Schroder, *et al.*, "The Peptides", vol. 1, Academic Press (New York), 1965 for solution synthesis.

In general, these methods comprise the sequential addition of one or more amino acids, or suitably protected amino acids, to a growing peptide chain. Normally, a suitable protecting group protects either the amino or carboxyl group of the first amino acid. The protected or derivatized amino acid can then be either attached to an inert solid support or utilized in solution by adding the

5   next amino acid in the sequence having the complementary (amino or carboxyl) group suitably protected, under conditions suitable for forming the amide linkage. The protecting group is then removed from this newly added amino acid residue and the next amino acid (suitably protected) is then added, and so forth. After all the desired amino acids have been linked in the proper sequence, any remaining protecting groups (and any solid support) are removed sequentially or

10   concurrently, to afford the final peptide. The protecting groups are removed, as desired, according to known methods depending on the particular protecting group utilized. For example, the protecting group may be removed by reduction with hydrogen and palladium on charcoal, sodium in liquid ammonia, etc.; hydrolysis with trifluoroacetic acid, hydrofluoric acid, and the like.

15   ## Lipophilic Sensitizers for Producing Active Species

A sensitizer is a chemical compound that can be induced to produce an active species that is capable of cleaving a cleavable linkage, preferably by oxidation. Preferably, the active species is a chemical species that exhibits short-lived activity so that its cleavage-inducing effects are only in the proximity of the site of its generation. Either the active species is inherently short lived, so that

20   it will not create significant background because beyond the proximity of its creation, or a scavenger is employed that efficiently scavenges the active species, so that it is not available to react with cleavable linkages beyond a short distance from the site of its generation. Illustrative active species include singlet oxygen, hydrogen peroxide, NADH, and hydroxyl radicals, phenoxy radical, superoxide, and the like. Illustrative quenchers for active species that cause oxidation

25   include polyenes, carotenoids, vitamin E, vitamin C, amino acid-pyrrole N-conjugates of tyrosine, histidine, and glutathione, and the like, e.g. Beutner et al, Meth. Enzymol., 319: 226-241 (2000).

In accordance with the invention, membranes in a sample are combined with a lipophilic sensitizer to form a sensitizer-treated membrane. A lipophilic sensitizer is formed by derivatizing a sensitizer either directly or indirectly with a lipophilic moiety that allows a sensitizer to be stably

30   anchored in a biological membrane. A lipophilic compound with a reactive functionality may be reacted with a complementary functionality on a sensitizer or a cross-linking agent to produce a sensitizer having a covalently attached lipophilic group for anchoring it in a membrane. Alternatively, a capture moiety, such as biotin, having a lipophilic group may be anchored in a membrane followed by conjugation to avidin or streptavidin and finally by attachment of a

35   biotinylated sensitizer via the avidin or streptavidin.

An important consideration for a lipophilic sensitizer and cleavable linkage is that they not be so far from one another that when a binding compound is bound to a membrane-associated

19

analyte the active species generated by the sensitizer diffuses and loses its activity before it can interact with the cleavable linkage. Accordingly, during a cleavage step, a sensitizer preferably is within 1000 nm, preferably 20-100 nm of a bound cleavage-inducing moiety. This effective range of a cleavage-inducing moiety is referred to herein as its "effective proximity."

5   Sensitizers for generating active species include enzymes, such as oxidases, such as glucose oxidase, xanthene oxidase, D-amino acid oxidase, NADH-FMN oxidoreductase, galactose oxidase, glyceryl phosphate oxidase, sarcosine oxidase, choline oxidase and alcohol oxidase, that produce hydrogen peroxide, horse radish peroxidase, that produces hydroxyl radical, various dehydrogenases that produce NADH or NADPH, urease that produces ammonia to create a high

10 local pH.

   A preferred sensitizer for use with the invention is a photosensitizer that generates singlet oxygen from molecular oxygen in response to photoexcitation. As used herein, "photosensitizer" refers to a light-adsorbing molecule that when activated by light converts molecular oxygen into singlet oxygen. Suitable photosensitizers having lipophilic moieties are disclosed in the following

15 references: Young et al, U.S. patent 6,375,930; and Young et al, U.S. patent application 2002/0006378, which are incorporated by reference. Additional photosensitizers that may be derivatized with lipophilic groups or capture moieties, such as biotin, and used with the invention are disclosed in the following references: Sessler et al, U.S. patent 5,292,414; Masuya et al, U.S. patent 5,344,928; McCapra, U.S. patent 5,705,622; Levy et al, 4,883,790; Meunier et al, U.S.

20 patent 5,141,911; and the like, which are incorporated by reference. The following references disclose the use of conjugates between biotin and lipophilic moieties to anchor biotinylated molecules to membranes via an avidin or streptavidin: Plant et al, Anal. Biochem., 176: 420-426 (1989); Bayer et al, Biochim. Biophys. Acta, 550: 464-473 (1979); Ramirez et al, J. Chromatogr. A, 971: 117-127 (2002); and the like, which are incorporated by reference.

25   Photosensitizers include dyes and aromatic compounds, and are usually compounds comprised of covalently bonded atoms, usually with multiple conjugated double or triple bonds. The compounds typically absorb light in the wavelength range of about 200 to about 1,100 nm, usually, about 300 to about 1,000 nm, preferably, about 450 to about 950 nm, with an extinction coefficient at its absorbance maximum greater than about 500 M^{-1} cm^{-1}, preferably, about 5,000 M^{-1}

30 cm^{-1}, more preferably, about 50,000 M^{-1} cm^{-1}, at the excitation wavelength. The lifetime of an excited state produced following absorption of light in the absence of oxygen will usually be at least about 100 nanoseconds, preferably, at least about 1 millisecond. In general, the lifetime must be sufficiently long to permit cleavage of a linkage in a reagent in accordance with the present invention. Such a reagent is normally present at concentrations as discussed below. The

35 photosensitizer excited state usually has a different spin quantum number (S) than its ground state and is usually a triplet (S=1) when the ground state, as is usually the case, is a singlet (S=0). Preferably, the photosensitizer has a high intersystem crossing yield. That is, photoexcitation of a

20

photosensitizer usually produces a triplet state with an efficiency of at least about 10%, desirably at least about 40%, preferably greater than about 80%.

Photosensitizers chosen are relatively photostable and, preferably, do not react efficiently with singlet oxygen. Several structural features are present in most useful photosensitizers. Most photosensitizers have at least one and frequently three or more conjugated double or triple bonds held in a rigid, frequently aromatic structure. They will frequently contain at least one group that accelerates intersystem crossing such as a carbonyl or imine group or a heavy atom selected from rows 3-6 of the periodic table, especially iodine or bromine, or they may have extended aromatic structures.

A large variety of light sources are available to photo-activate photosensitizers to generate singlet oxygen. Both polychromatic and monchromatic sources may be used as long as the source is sufficiently intense to produce enough singlet oxygen in a practical time duration. The length of the irradiation is dependent on the nature of the photosensitizer, the nature of the cleavable linkage, the power of the source of irradiation, and its distance from the sample, and so forth. In general, the period for irradiation may be less than about a microsecond to as long as about 10 minutes, usually in the range of about one millisecond to about 60 seconds. The intensity and length of irradiation should be sufficient to excite at least about 0.1% of the photosensitizer molecules, usually at least about 30% of the photosensitizer molecules and preferably, substantially all of the photosensitizer molecules. Exemplary light sources include, by way of illustration and not limitation, lasers such as, e.g., helium-neon lasers, argon lasers, YAG lasers, He/Cd lasers, and ruby lasers; photodiodes; mercury, sodium and xenon vapor lamps; incandescent lamps such as, e.g., tungsten and tungsten/halogen; flashlamps; and the like.

Examples of photosensitizers that may be utilized in the present invention are those that have the above properties and are enumerated in the following references: Turro, Modern Molecular Photochemistry (cited above); Singh and Ullman, U.S. patent 5,536,834; Li et al, U.S. patent 5,763,602; Ullman, et al., Proc. Natl. Acad. Sci. USA 91, 5426-5430 (1994); Strong et al, Ann. New York Acad. Sci., 745: 297-320 (1994); Martin et al, Methods Enzymol., 186: 635-645 (1990);Yarmush et al, Crit. Rev. Therapeutic Drug Carrier Syst., 10: 197-252 (1993); Pease et al, U.S. patent 5,709,994; Ullman et al, U.S. patent 5,340,716; Ullman et al, U.S. patent 6,251,581; McCapra, U.S. patent 5,516,636; Wohrle, Chimia, 45: 307-310 (1991); Thetford, European patent publ. 0484027; Sessler et al, SPIE, 1426: 318-329 (1991); Madison et al, Brain Research, 522: 90-98 (1990); Polo et al, Inorganica Chimica Acta, 192: 1-3 (1992); Demas et al, J. Macromol. Sci., A25: 1189-1214 (1988); and the like. In one embodiment, photosensitizers used in the invention are porphyrins, e.g. as described in Roelant, U.S. patent 6,001,573, which is incorporated by reference. Many porphyrins suitable for use with the invention are available commercially, e.g. Frontier Scientific, Inc. (Logan, Utah); Molecular Probes, Inc. (Eugene, Oregon); and the like. Exemplary photosensitizers are listed in Table 1b.

21

Exemplary Photosensitizers

Hypocrellin A	Tetraphenylporphyrin
Hypocrellin B	Halogenated derivatives of rhodamine dyes
Hypericin	metallo-Porphyrins
Halogenated derivatives of fluorescein dyes	Phthalocyanines
Rose bengal	Naphthalocyanines
Merocyanine 540	Texaphyrin-type macrocycles
Methylene blue	Hematophorphyrin
9-Thioxanthone	9,10-Dibromoanthracene
Chlorophylls	Benzophenone
Phenaleone	Chlorin e6
Protoporphyrin	Perylene
Benzoporphryin A monacid	Benzoporphryin B monacid

Treating Membranes with Lipophilic Sensitizers

Lipophilic sensitizers can be incorporated into lipid membranes in an orientation and manner similar to that of phospholipids where the hydrophobic moiety comprising of the hydrocarbon chains can orient inward and the more hydrophilic entities can orient outwards. Thus, as in the usual cellular membranes, the hydrocarbon portion of the lipophilic sensitizer can be incorporated into the lipid environment whereas the hydrophilic sensitizer portion can be exposed to the aqueous interface at the membrane surface.

Where intact cellular structures are required, the methods used to incorporate lipophilic sensitizers into the cells preferably cause minimal disruption of the cell and of the integrity of membranes. In addition, the cells can be fixed and treated with routine histochemical or cytochemical procedures, where the procedure preferably does not affect the incorporation.

The membranes can be labeled with lipophilic sensitizers according to the method described in Barak and Webb (1981) J. Cell Biol. 90:595-604. Typically, the membrane, such as the intact cell, is contacted with the compounds of the invention, preferably in an aqueous media. The aqueous media can be water, water and organic solvent, such as DMSO, DMF, DMA, or a mixture thereof, and can contain buffers such as phosphate, acetate, tris, and the like. The membranes and lipophilic sensitizer are contacted for between 1 min. to about 1 week, preferably about 1 h to 76 h, more preferably about 2h to about 48 h, or any integer in between. The formulations may additionally be subjected to chemical or mechanical treatment, such as the addition of a surfactant (Tween 80, for example), shaking, stirring, electroporation, and the like. Alternatively, the formulation can be heated to about 30 °C to 50 °C, preferably about 35 °C to about 40 °C, until labeling is achieved. After labeling, the unbound components can be removed by washing, or by centrifugation, for example, and the sensitizer-labeled cells or membranes isolated.

Alternatively, a capture moiety, such as biotin, having a lipophilic group may be anchored in a membrane followed by conjugation to avidin or streptavidin and finally by attachment of a biotinylated sensitizer via the avidin or streptavidin, as illustrated in Figure 1D. Cells (1301) are combined with biotin having a lipophilic moiety (1304) (referred to below as biotin-G) to form a

5   population of cells (1306) having membranes containing free biotin. To this population is added avidin or streptavidin (1310) to form biotin-avidin or biotin-streptavidin complexes (1312) on the cell surfaces. These cells are then combined (1313) with biotinylated photosensitizers to form complexes (1312) on the cell surfaces that comprise sensitizer-treated membranes.

10   Separation of Released Molecular Tags

As mentioned above, molecular tags are designed for separation by a separation technique that can distinguish molecular tags based on one or more physical, chemical, and/or optical characteristics. Preferably, such separation technique is capable of providing quantitative information as well as qualitative information about the presence or absence of molecular tags (and

15   therefore, corresponding analytes). In one aspect, a liquid phase separation technique is employed so that a solution, e.g. buffer solution, reaction solvent, or the like, containing a mixture of molecular tags is processed to bring about separation of individual kinds of molecular tags. Usually, such separation is accompanied by the differential movement of molecular tags from such a starting mixture along a path until discernable peaks or bands form that correspond to regions of

20   increased concentration of the respective molecular tags. Such a path may be defined by a fluid flow, electric field, magnetic field, or the like. The selection of a particular separation technique depends on several factors including the expense and convenience of using the technique, the resolving power of the technique given the chemical nature of the molecular tags, the number of molecular tags to be separated, the type of detection mode employed, and the like. Preferably,

25   molecular tags are electrophoretically or chromatographically separated.

A. Electrophoretic Separation

Methods for electrophoresis of are well known and there is abundant guidance for one of ordinary skill in the art to make design choices for forming and separating particular pluralities of

30   molecular tags. The following are exemplary references on electrophoresis: Krylov et al, Anal. Chem., 72: 111R-128R (2000); P.D. Grossman and J.C. Colburn, Capillary Electrophoresis: Theory and Practice, Academic Press, Inc., NY (1992); U.S. Patents 5,374,527; 5,624,800; 5,552,028; ABI PRISM 377 DNA Sequencer User's Manual, Rev. A, January 1995, Chapter 2 (Applied Biosystems, Foster City, CA); and the like. In one aspect, molecular tags are separated by

35   capillary electrophoresis. Design choices within the purview of those of ordinary skill include but are not limited to selection of instrumentation from several commercially available models, selection of operating conditions including separation media type and concentration, pH, desired

separation time, temperature, voltage, capillary type and dimensions, detection mode, the number of molecular tags to be separated, and the like.

In one aspect of the invention, during or after electrophoretic separation, the molecular tags are detected or identified by recording fluorescence signals and migration times (or migration distances) of the separated compounds, or by constructing a chart of relative fluorescent and order of migration of the molecular tags (*e.g.*, as an electropherogram). To perform such detection, the molecular tags can be illuminated by standard means, *e.g.* a high intensity mercury vapor lamp, a laser, or the like. Typically, the molecular tags are illuminated by laser light generated by a He-Ne gas laser or a solid-state diode laser. The fluorescence signals can then be detected by a light-sensitive detector, *e.g.*, a photomultiplier tube, a charged-coupled device, or the like. Exemplary electrophoresis detection systems are described elsewhere, *e.g.*, U.S. Patent Nos. 5,543,026; 5,274,240; 4,879,012; 5,091,652; 6,142,162; or the like. In another aspect, molecular tags may be detected electrochemically detected, e.g. as described in U.S. Patent No. 6,045,676.

Electrophoretic separation involves the migration and separation of molecules in an electric field based on differences in mobility. Various forms of electrophoretic separation include, by way of example and not limitation, free zone electrophoresis, gel electrophoresis, isoelectric focusing, isotachophoresis, capillary electrochromatography, and micellar electrokinetic chromatography. Capillary electrophoresis involves electroseparation, preferably by electrokinetic flow, including electrophoretic, dielectrophoretic and/or electroosmotic flow, conducted in a tube or channel of from about 1 to about 200 micrometers, usually, from about 10 to about 100 micrometers cross-sectional dimensions. The capillary may be a long independent capillary tube or a channel in a wafer or film comprised of silicon, quartz, glass or plastic.

In capillary electroseparation, an aliquot of the reaction mixture containing the molecular tags is subjected to electroseparation by introducing the aliquot into an electroseparation channel that may be part of, or linked to, a capillary device in which the amplification and other reactions are performed. An electric potential is then applied to the electrically conductive medium contained within the channel to effectuate migration of the components within the combination. Generally, the electric potential applied is sufficient to achieve electroseparation of the desired components according to practices well known in the art. One skilled in the art will be capable of determining the suitable electric potentials for a given set of reagents used in the present invention and/or the nature of the cleaved labels, the nature of the reaction medium and so forth. The parameters for the electroseparation including those for the medium and the electric potential are usually optimized to achieve maximum separation of the desired components. This may be achieved empirically and is well within the purview of the skilled artisan.

24

Detection may be by any of the known methods associated with the analysis of capillary electrophoresis columns including the methods shown in U.S. Patent Nos. 5,560,811 (column 11, lines 19-30), 4,675,300, 4,274,240 and 5,324,401, the relevant disclosures of which are incorporated herein by reference. Those skilled in the electrophoresis arts will recognize a wide

5   range of electric potentials or field strengths may be used, for example, fields of 10 to 1000 V/cm are used with about 200 to about 600 V/cm being more typical. The upper voltage limit for commercial systems is about 30 kV, with a capillary length of about 40 to about 60 cm, giving a maximum field of about 600 V/cm. For DNA, typically the capillary is coated to reduce electroosmotic flow, and the injection end of the capillary is maintained at a negative potential.

10   For ease of detection, the entire apparatus may be fabricated from a plastic material that is optically transparent, which generally allows light of wavelengths ranging from about 180 to about 1500 nm, usually about 220 to about 800 nm, more usually about 450 to about 700 nm, to have low transmission losses. Suitable materials include fused silica, plastics, quartz, glass, and so forth.

15   B. Chromatographic Separation

In one aspect of the invention, pluralities of molecular tags are designed for separation by chromatography based on one or more physical characteristics that include but are not limited to molecular weight, shape, solubility, pKa, hydrophobicity, charge, polarity, or the like. A chromatographic separation technique is selected based on parameters such as column type, solid

20   phase, mobile phase, and the like, followed by selection of a plurality of molecular tags that may be separated to form distinct peaks or bands in a single operation. Several factors determine which HPLC technique is selected for use in the invention, including the number of molecular tags to be detected (i.e. the size of the plurality), the estimated quantities of each molecular tag that will be generated in the assays, the availability and ease of synthesizing molecular tags that are candidates

25   for a set to be used in multiplexed assays, the detection modality employed, and the availability, robustness, cost, and ease of operation of HPLC instrumentation, columns, and solvents. Generally, columns and techniques are favored that are suitable for analyzing limited amounts of sample and that provide the highest resolution separations. Guidance for making such selections can be found in the literature, e.g. Snyder et al, Practical HPLC Method Development, (John Wiley

30   & Sons, New York, 1988); Millner, "High Resolution Chromatography: A Practical Approach", Oxford University Press, New York (1999), Chi-San Wu, "Column Handbook for Size Exclusion Chromatography", Academic Press, San Diego (1999), and Oliver, "HPLC of Macromolecules: A Practical Approach, Oxford University Press", Oxford, England (1989). In particular, procedures are available for systematic development and optimization of chromatographic separations given

35   conditions, such as column type, solid phase, and the like, e.g. Haber et al, J. Chromatogr. Sci., 38: 386-392 (2000); Outinen et al, Eur. J. Pharm. Sci., 6: 197-205 (1998); Lewis et al, J. Chromatogr., 592: 183-195 and 197- 208 (1992); and the like.

25

In one aspect, initial selections of molecular tag candidates are governed by the physiochemical properties of molecules typically separated by the selected column and stationary phase. The initial selections are then improved empirically by following conventional optimization procedure, as described in the above reference, and by substituting more suitable candidate

5 molecular tags for the separation objectives of a particular embodiment. In one aspect, separation objectives of the invention include (i) separation of the molecular tags of a plurality into distinguishable peaks or bands in a separation time of less than 60 minutes, and more preferably in less than 40 minutes, and still more preferably in a range of between 10 to 40 minutes, (ii) the formation of peaks or bands such that any pair has a resolution of at least 1.0, more preferably at

10 least 1.25, and still more preferably, at least 1.50, (iii) column pressure during separation of less than 150 bar, (iv) separation temperature in the range of from 25°C to 90°C, preferably in the range of from 35°C to 80°C, and (v) the plurality of distinguishable peaks is in the range of from 5 to 30 and all of the peaks in the same chromatogram. As used herein, "resolution" in reference to two peaks or bands is the distance between the two peak or band centers divided by the average base

15 width of the peaks, e.g. Snyder et al (cited above).

A chromatographic method is used to separate molecular tags based on their chromatographic properties. A chromatographic property can be, for example, a retention time of a molecular tag on a specific chromatographic medium under defined conditions, or a specific condition under which a molecular tag is eluted from a specific chromatographic medium. A

20 chromatographic property of a molecular tag can also be an order of elution, or pattern of elution, of a molecular tag contained in a group or set of molecular tags being chromatographically separated using a specific chromatographic medium under defined conditions. A chromatographic property of a molecular tag is determined by the physical properties of the molecular tag and its interactions with a chromatographic medium and mobile phase. Defined conditions for

25 chromatography include particular mobile phase solutions, column geometry, including column diameter and length, pH, flow rate, pressure and temperature of column operation, and other parameters that can be varied to obtain the desired separation of molecular tags. A molecular tag, or chromatographic property of a molecular tag, can be detected using a variety of chromatography methods.

30 Although standard liquid chromatography methods can be used to separate molecular tags, high pressure (or performance) liquid chromatography (HPLC) provides the advantages of high resolution, increased speed of analysis, greater reproducibility, and ease of automation of instrument operation and data analysis. HPLC methods also allow separation of molecular tags based on a variety of physiochemical properties. Molecular tags having similar properties can be

35 used together in the same experiment since HPLC can be used to differentiate between closely related tags. The high degree of resolution achieved using HPLC methods allows the use of large

26

sets of tagged probes because the resulting molecular tags can be distinguished from each other. The ability to detect large sets of tagged probes is an advantage when performing multiplexed detection of target nucleic acids and target analytes. As used herein, "HPLC" refers to a liquid phase chromatographic separation that (i) employs a rigid cylindrical separation column having a

5 length of up to 300 mm and an inside diameter of up to 5 mm, (ii) has a solid phase comprising rigid spherical particles (e.g. silica, alumina, or the like) having the same diameter of up to 5 μm packed into the separation column, (iii) takes place at a temperature in the range of from 35°C to 80°C and at column pressure up to 150 bars, and (iv) employs a flow rate in the range of from 1 μL/min to 4 mL/min. Solid phase particles for use in HPLC are further characterized in (i) having

10 a narrow size distribution about the mean particle diameter, with substantially all particle diameters being within 10% of the mean, (ii) having the same pore size in the range of from 70 to 300 angstroms, (iii) having a surface area in the range of from 50 to 250 m^2/g, and (iv) having a bonding phase density (i.e. the number of retention ligands per unit area) in the range of from 1 to 5 per nm^2.

15   Sets of molecular tags detected in a single experiment generally are a group of chemically related molecules that differ by mass, charge, mass-charge ratio, detectable tag, such as differing fluorophores or isotopic labels, or other unique characteristic. Therefore, both the chemical nature of the molecular tag and the particular differences among molecular tags in a group of molecular tags can be considered when selecting a suitable chromatographic medium for separating molecular

20 tags in a sample.

Separation of molecular tags by liquid chromatography can be based on physical characteristics of molecular tags such as charge, size and hydrophobicity of molecular tags, or functional characteristics such as the ability of molecular tags to bind to molecules such as dyes, lectins, drugs, peptides and other ligands on an affinity matrix. A wide variety of chromatographic

25 media are suitable for separation of molecular tag based on charge, size, hydrophobicity and other chromatographic properties of molecular tags. Selection of a particular chromatographic medium will depend upon the properties of molecular tags employed.

Separation of molecular tags based on charge can be performed by ion exchange chromatography. Methods for separating peptides, proteins, oligonucleotides, and nucleic acids are

30 well known to those skilled in the art and are described, for example, in Millner, supra (1999). In this technique, separation is based on the exchange of ions (anions or cations) between the mobile phase and ionic sites on the stationary phase. Charged chemical species are covalently bound to the surface of the stationary phase to prepare an ion exchange resin. The mobile phase contains a large number of counterions that are opposite in charge to the resin ionic group to form an ion-pair.

35 A molecular tag having the same ionic charge as the counterion will be in equilibrium with the counterion. The molecular tag ion can exchange with the counter ion to pair with the covalently

27

attached charge on the support. When the molecular tag ion is paired with the charged group on the support, it does not move through the column. Molecular tag ion retention is based on the affinity of different ions on the support and other solution parameters including counterion type, ionic strength and pH.

5   Ion exchange media fall into two classes that include strong ion exchangers and weak ion exchangers. The charge of weak ion exchangers varies with pH of the mobile phase, while the charge of strong ion exchangers is essentially independent of pH. In most cases, it is advantageous to select a strong exchanger to separate molecular tags, but when molecular tags bind very tightly to strong exchangers, a weak exchanger is advantageous to allow maximum recovery of molecular

10   tags.

       Ion exchange media useful for separating molecular tags include both anion or cation exchangers. The choice of whether to use an anion or cation exchanger to separate molecular tags will therefore depend on the charge of the molecular tags at the pH of the chromatographic step. The choice of the pH for the separation can be selected by determining the isolelectric point (pI) of

15   the molecular tag, or the average isoelectric point of a group of molecular tags, and generally using one pH unit above the pI for anion exchange or one pH unit below the pI for cation exchange.

       Cation exchange resins have anionic functional groups such as $-SO_3-$, $-OPO_3-$ and $-COO-$ and anion exchange matrices usually contain the cationic tertiary and quaternary ammonium groups, with general formulae $-NHR_2+$ and $-NR_3+$. Exemplary ion exchange chromatography

20   media for separating molecular tags that are peptides, polypeptides, nucleic acids and chemical compounds include strong and weak anion and cation exchange resins having functional groups such as sulfonic acid, quaternary amine and tertiary amine, commonly known as S, Q, and DEAE resins, respectively.

       Separation of molecular tags that are smaller molecules, such as chemical compounds, for

25   example alkylenes and aralkylenes, can be performed using small pore size resins, whereas wide-pore resins generally are used for separating molecular tags that are peptides, polypeptides and nucleic acid molecules.

       Separation of molecular tags based on hydrophobic interactions can be performed by hydrophobic interaction chromatography and closely related reversed-phase chromatography

30   methods. Hydrophobic interaction chromatography (HIC) has generally been most useful for separating small molecules and peptides, while reversed phase chromatography has been more widely applicable to larger molecules, such as polypeptides and nucleic acids. HIC employs a chemically bonded hydrophobic stationary phase, with the mobile phase being more polar than the stationary phase. The basis of HIC is the interaction between hydrophobic parts of molecular tags

35   and a hydrophobic matrix. HIC can be used to separate a variety of types of molecular tags,

28

including organic molecules, oligonucleotides and peptides. Exemplary HIC chromatography media for separating molecular tags that are oligonucleotides, peptides or chemical compounds, include phenyl, butyl or octyl hydrophobic ligands coupled to a sepharose matrix and ether, isopropyl or hydrophobic ligands coupled to a polystyrene/divinylbenzene matrix.

5   Reverse phase chromatography is a type of chromatography in which the chemically bonded phase is hydrophobic (nonpolar) than the mobile phase. This is "reversed" from normal phase chromatography, in which the stationary phase is hydrophilic (polar), and the starting mobile phase is more nonpolar than the stationary phase. Mobile phase gradients that increase in concentration of an organic modifier (usually acetonitrile or methanol) are commonly used in
10   reverse phase HPLC. These gradients elute solute molecules in order of increasing hydrophobicity. Exemplary mobile phases for use with the invention to separate water soluble molecular tags include but are not limited to water, nitromethane, methanol, dimethyl sulfoxide, dimethylformamide, acetonitrile, acetic acid, methoxyethanol, benzyl alcohol, acetone, and the like. The mobile phases may be used isocratically or they may be combined and delivered to a
15   column in continuously varying proportions. In the latter case, usually two solvents are combined in proportions that vary linearly over time, i.e. gradient delivery.

Various mobile phase additives can be used to provide different selectivity to improve separation of molecular tags. For example, ion pairing reagents may be used in reverse phase HPLC methods. Exemplary ion pairing reagents include trifluoroacetic acid (TFA), which is an
20   anionic ion-pairing reagent, and tetrabutylammonium phosphate, which is a cationic ion pairing reagent.

Reverse phase HPLC can be used to separate a variety of types of molecular tags, including organic molecules, oligonucleotides, peptides and polypeptides. Reversed phase HPLC is particularly useful for separating peptide or polypeptide molecular tags that are closely related to
25   each other. Exemplary reversed phase chromatography media for separating molecular tags include particles, e.g. silica or alumina, having bonded to their surfaces retention ligands, such as phenyl groups, cyano groups, or aliphatic groups selected from the group including C_8 through C_{18}. Preferably, the particles have a pore size in the range of from 80 to 300 angstroms.

Exemplary reversed phase chromatography media for separating molecular tags that are
30   peptides, include particles having aliphatic retention ligands in the range of from C_8 to C_{18} bonded to their surfaces and having a pore size of between 60 and 80 angstroms. Commercial preparations useful for separating molecular tags include, for example, Apex WP Octadecyl C_{18}, Octyl C_8, Butyl C_4 and Phenyl, Aquaprep RP-3000 C_4 and C_8, Bakerbond WP Octadecyl C_{18}, Octyl C_8, Butyl C_4 and Diphenyl.

29

When reverse phase or ion-pair HPLC methods are insufficient to provide adequate separation of all molecular tags, switching to normal phase HPLC may be helpful, because different retention processes provide different selectivity effects. In contrast to the conditions used for reversed phase chromatography, normal phase chromatography involves using a stationary

5   phase is hydrophilic (polar), and the starting mobile phase is more non-polar than the stationary phase. Sample retention is controlled by adsorption to the stationary phase, and molecules must displace solvent molecules from the stationary phase. Normal phase chromatography can be used to separate molecular tags having a variety of physicochemical properties.

Mixed mode chromatography also can be used to separate molecular tags, and is

10   particularly useful for separating oligonucleotide reporter tags. Mixed mode chromatography takes advantage of both hydrophobic and electrostatic interactions between the molecular tags to be separated and the stationary phase. Exemplary mixed mode column packing materials include NACS-12, derivatized aminopropyl silica particles with alkyl and aryl residues.

Prior to separation by HPLC, a sample can be fractionated or subjected to a pre-separation

15   step, for example, to remove particulate matter or molecules other than reporter tags. In addition to standard biochemical methods for fractionating samples, such as centrifugation, precipitation, filtration and extraction, a variety of HPLC pre-columns or guard columns can be used for this purpose.

Separated molecular tags can be detected using a variety of analytical methods, including

20   detection of intrinsic properties of molecular tags, such as absorbance, fluorescence or electrochemical properties, as well as detection of a detection group or moiety attached to a molecular tag. Although not required, a variety of detection groups or moieties can be attached to molecular tags to facilitate detection after chromatographic separation.

Detection methods for use with liquid chromatography are well known, commercially

25   available, and adaptable to automated and high-throughput sampling. The detection method selected for analysis of molecular tags will depend upon whether the molecular tags contain a detectable group or moiety, the type of detectable group used, and the physicochemical properties of the molecular tag and detectable group, if used. Detection methods based on fluorescence, electrolytic conductivity, refractive index, and evaporative light scattering can be used to detect

30   various types of molecular tags.

A variety of optical detectors can be used to detect a molecular tag separated by liquid chromatography. Methods for detecting nucleic acids, polypeptides, peptides, and other macromolecules and small molecules using ultraviolet (UV)/visible spectroscopic detectors are well known, making UV/visible detection the most widely used detection method for HPLC

analysis. Infrared spectrophotometers also can be used to detect macromolecules and small molecules when used with a mobile phase that is a transparent polar liquid.

Variable wavelength and diode-array detectors represent two commercially available types of UV/visible spectrophotometers. A useful feature of some variable wavelength UV detectors is

5   the ability to perform spectroscopic scanning and precise absorbance readings at a variety of wavelengths while the peak is passing through the flowcell. Diode array technology provides the additional advantage of allowing absorbance measurements at two or more wavelengths, which permits the calculation of ratios of such absorbance measurements. Such absorbance rationing at multiple wavelengths is particularly helpful in determining whether a peak represents one or more

10   than one molecular tag.

Fluorescence detectors can also be used to detect fluorescent molecular tags, such as those containing a fluorescent detection group and those that are intrinsically fluorescent. Typically, fluorescence sensitivity is relatively high, providing an advantage over other spectroscopic detection methods when molecular tags contain a fluorophore. Although molecular tags can have

15   detectable intrinsic fluorescence, when a molecular tag contains a suitable fluorescent detection group, it can be possible to detect a single molecular tag in a sample.

Electrochemical detection methods are also useful for detecting molecular tags separated by HPLC. Electrochemical detection is based on the measurement of current resulting from oxidation or reduction reaction of the molecular tags at a suitable electrode. Since the level of

20   current is directly proportional to molecular tag concentration, electrochemical detection can be used quantitatively, if desired.

Evaporative light scattering detection is based on the ability of particles to cause photon scattering when they traverse the path of a polychromatic beam of light. The liquid effluent from an HPLC is first nebulized and the resultant aerosol mist, containing the molecular tags, is directed

25   through a light beam. A signal is generated that is proportional to the amount of the molecular tag present in a sample, and is independent of the presence or absence of detectable groups such as chromophores, fluorophores or electroactive groups. Therefore, the presence of a detection group or moiety on a molecular tag is not required for evaporative light scattering detection.

Mass spectrometry methods also can be used to detect molecular tags separated by HPLC.

30   Mass spectrometers can resolve ions with small mass differences and measure the mass of ions with a high degree of accuracy and sensitivity. Mass spectrometry methods are well known in the art (see Burlingame et al. Anal. Chem. 70:647R-716R (1998); Kinter and Sherman, Protein Sequencing and Identification Using Tandem Mass Spectrometry Wiley-Interscience, New York (2000)).

31

Analysis of data obtained using any detection method, such as spectral deconvolution and quantitative analysis can be manual or computer-assisted, and can be performed using automated methods. A variety of computer programs can be used to determine peak integration, peak area, height and retention time. Such computer programs can be used for convenience to determine the presence of a molecular tag qualitatively or quantitatively. Computer programs for use with HPLC and corresponding detectors are well known to those skilled in the art and generally are provided with commercially available HPLC and detector systems.

The particular molecular tags contained in a sample can be determined, for example, by comparison with a database of known chromatographic properties of reference molecular tags, or by algorithmic methods such as chromatographic pattern matching, which allows the identification of components in a sample without the need to integrate the peaks individually. The identities of molecular tags in a sample can be determined by a combination of methods when large numbers of molecular tags are simultaneously identified, if desired.

A variety of commercially available systems are well-suited for high throughput analysis of molecular tags. Those skilled in the art can determine appropriate equipment, such as automated sample preparation systems and autoinjection systems, useful for automating HPLC analysis of molecular tags. Automated methods can be used for high-throughput analysis of molecular tags, for example, when a large number of samples are being processes or for multiplexed application of the methods of the invention for detecting target analytes. An exemplary HPLC instrumentation system suitable for use with the present invention is the Agilent 1100 Series HPLC system (Agilent Technologies, Palo Alto, CA).

Those skilled in the art will be aware of quality control measures useful for obtaining reliable analysis of molecular tags, particular when analysis is performed in a high-throughput format. Such quality control measures include the use of external and internal reference standards, analysis of chromatograph peak shape, assessment of instrument performance, validation of the experimental method, for example, by determining a range of linearity, recovery of sample, solution stability of sample, and accuracy of measurement.

In another aspect of the invention, molecular tags are separated by capillary electrochromatography (CEC). In CEC, the liquid phase is driven by electroosmotic flow through a capillary-sized column, e.g. with inside diameters in the range of from 30 to 100 μm. CEC is disclosed in Svec, Adv. Biochem. Eng. Biotechnol. 76: 1-47 (2002); Vanhoenacker et al, Electrophoresis, 22: 4064-4103 (2001); and like references. CEC column may used the same solid phase materials as used in conventional reverse phase HPLC and additionally may use so-called "monolithic" non-particular packings. In some forms of CEC, pressure as well as electroosmosis drives a sample-containing solvent through a column.

32

Exemplary Synthetic Approaches for Molecular Tags

One exemplary synthetic approach is outlined in Figure 1. Starting with commercially available 6-carboxy fluorescein, the phenolic hydroxyl groups are protected using an anhydride. Isobutyric anhydride in pyridine was employed but other variants are equally suitable. It is

5 important to note the significance of choosing an ester functionality as the protecting group. This species remains intact throughout the phosphoramidite monomer synthesis as well as during oligonucleotide construction. These groups are not removed until the synthesized oligonucleotide is deprotected using ammonia. After protection the crude material is then activated *in situ* via formation of an N-hydroxysuccinimide ester (NHS-ester) using DCC as a coupling agent. The

10 DCU by product is filtered away and an amino alcohol is added. Many amino alcohols are commercially available some of which are derived from reduction of amino acids. When the amino alcohol is of the form "$H_2N-(CH_2)_n-OH$," n is in the range of from 2 to 12, and more preferably, from 2 to 6. Only the amine is reactive enough to displace N-hydroxysuccinimide. Upon standard extractive workup, a 95% yield of product is obtained. This material is phosphitylated to generate

15 the phosphoramidite monomer. For the synthesis of additional molecular tags, a symmetrical bis-amino alcohol linker is used as the amino alcohol (Figure 2). As such, the second amine is then coupled with a multitude of carboxylic acid derivatives (exemplified by several possible benzoic acid derivatives shown in Figure 3 prior to the phosphitylation reaction.

Alternatively, molecular tags may be made by an alternative strategy that uses 5-
20 aminofluorescein as starting material (Figure 4). Addition of 5-aminofluorescein to a great excess of a diacid dichloride in a large volume of solvent allows for the predominant formation of the monoacylated product over dimer formation. The phenolic groups are not reactive under these conditions. Aqueous workup converts the terminal acid chloride to a carboxylic acid. This product is analogous to 6-carboxyfluorescein, and using the same series of steps is converted to its

25 protected phosphoramidite monomer. There are many commercially available diacid dichlorides and diacids, which can be converted to diacid dichlorides using $SOCl_2$ or acetyl chloride. There are many commercial diacid dichlorides and amino alcohols (Figure 5). These synthetic approaches are ideally suited for combinatorial chemistry.

The molecular tags constructed with the schemes of Figures 1, 2, and 4 are further reacted
30 either before or after phosphitylation to attach a cleavable linkage, *e.g.* using chemistry as described below.

The molecular tag may be assembled having an appropriate functionality at one end for linking to the polypeptide-binding moieties. A variety of functionalities can be employed. Thus, the functionalities normally present in a peptide, such as carboxy, amino, hydroxy and thiol may be

35 the targets of a reactive functionality for forming a covalent bond. The molecular tag is linked in accordance with the chemistry of the linking group and the availability of functionalities on the polypeptide-binding moiety. For example, as discussed above for antibodies, and fragments

33

thereof such as Fab' fragments, specific for a polypeptide, a thiol group will be available for using an active olefin, e.g., maleimide, for thioether formation. Where lysines are available, one may use activated esters capable of reacting in water, such as nitrophenyl esters or pentafluorophenyl esters, or mixed anhydrides as with carbodiimide and half-ester carbonic acid. There is ample chemistry

5   for conjugation in the literature, so that for each specific situation, there is ample precedent in the literature for the conjugation.

In an illustrative synthesis a diol is employed. Examples of such diols include an alkylene diol, polyalkylene diol, with alkylene of from 2 to 3 carbon atoms, alkylene amine or poly(alkylene amine) diol, where the alkylenes are of from 2 to 3 carbon atoms and the nitrogens are substituted,

10   for example, with blocking groups or alkyl groups of from 1 − 6 carbon atoms, where one diol is blocked with a conventional protecting group, such as a dimethyltrityl group. This group can serve as the mass-modifying region and with the amino groups as the charge-modifying region as well. If desired, the mass modifier can be assembled by using building blocks that are joined through phosphoramidite chemistry. In this way the charge modifier can be interspersed between the mass

15   modifier. For example, a series of polyethylene oxide molecules having 1, 2, 3, n units may be prepared. To introduce a number of negative charges, a small polyethylene oxide unit may be employed. The mass and charge-modifying region may be built up by having a plurality of the polyethylene oxide units joined by phosphate units. Alternatively, by employing a large spacer, fewer phosphate groups would be present, so that without large mass differences, large differences

20   in mass-to- charge ratios may be realized.

The chemistry that is employed is the conventional chemistry used in oligonucleotide synthesis, where building blocks other than nucleotides are used, but the reaction is the conventional phosphoramidite chemistry and the blocking group is the conventional dimethoxytrityl group. Of course, other chemistries compatible with automated synthesizers can

25   also be used. However, it is desirable to minimize the complexity of the process.

As mentioned above, in one embodiment the hub nucleus is a hydrophilic polymer, generally, an addition or condensation polymer with multiple functionality to permit the attachment of multiple moieties. One class of polymers that is useful for the reagents of the present invention comprises the polysaccharide polymers such as dextrans, sepharose, polyribose, polyxylose, and

30   the like. For example, the hub may be dextran to which multiple molecular tags may be attached in a cleavable manner consistent with the present invention. A few of the aldehyde moieties of the dextran remain and may be used to attach the dextran molecules to amine groups on an oligonucleotide by reductive amination. In another example using dextran as the hub nucleus, the dextran may be capped with succinic anhydride and the resulting material may be linked to amine-

35   containing oligonucleotides by means of amide formation.

Besides the nature of the linker and mobility-modifying moiety, as already indicated, diversity can be achieved by the chemical and optical characteristics of the fluorescer, the use of

34

energy transfer complexes, variation in the chemical nature of the linker, which affects mobility, such as folding, interaction with the solvent and ions in the solvent, and the like. As already suggested, in one embodiment the linker is an oligomer, where the linker may be synthesized on a support or produced by cloning or expression in an appropriate host. Conveniently, polypeptides

5   can be produced where there is only one cysteine or serine/threonine/tyrosine, aspartic/glutamic acid, or lysine/arginine/histidine, other than an end group, so that there is a unique functionality, which may be differentially functionalized. By using protective groups, one can distinguish a side-chain functionality from a terminal amino acid functionality. Also, by appropriate design, one may provide for preferential reaction between the same functionalities present at different sites on the

10   linking group. Whether one uses synthesis or cloning for preparation of oligopeptides, will to a substantial degree depend on the length of the linker.

Methods of Using Binding Compositions of the Invention

In one aspect, the invention provides a method for detecting or measuring one or more

15   target analytes from biological sources. Conventional methodologies are employed to prepare samples for analysis. For example, for protein analytes guidance in sample preparation can be found in Scopes, Protein Purification, chapter 2 (Springer-Verlag, New York), where a range of procedures are disclosed for preparing protein extracts from different sources. Preparative techniques include mild cell lysis by osmotic disruption of cellular membranes, to enzymatic

20   digestion of connective tissue followed by osmotic-based lysis, to mechanical homogenization, to ultrasonication.

In some embodiments, a sample containing membrane-associated analytes of interest is treated with a lipophilic sensitizer as described above to form sensitizer-treated membranes. After such preparation, a reagent containing a plurality of binding compounds are added. The amounts

25   of binding compounds are usually determined empirically. Such components are combined under binding conditions, usually in an aqueous medium, generally at a pH in the range of about 5 to about 10, with buffer at a concentration in the range of about 10 to about 200 mM. These conditions are conventional, where conventional buffers may be used, such as phosphate, carbonate, HEPES, MOPS, Tris, borate, etc., as well as other conventional additives, such as salts,

30   stabilizers, organic solvents, etc. The aqueous medium may be solely water or may include from 0.01 to 80 or more volume percent of a co-solvent.

The reagents are incubated for a time and at a temperature that permit a substantial number of binding events to occur. The time for incubation after combination of the reagents varies depending on the (i) nature and expected concentration of the analyte being detected, (ii) the

35   mechanism by which the binding compounds for complexes with analytes, and (iii) the affinities of the specific reagents employed. Moderate temperatures are normally employed for the incubation

35

and usually constant temperature. Incubation temperatures will normally range from about 5° to 99°C, usually from about 15° to 85°C, more usually 35° to 75°C.

Generally, the concentrations of the various agents involved with an assay of the invention will vary with the concentration range of the individual analytes in the samples to be analyzed,
5   generally being in the range of about 10 nM to about 10 mM. Buffers will ordinarily be employed at a concentration in the range of about 10 to about 200 mM. The concentration of each analyte will generally be in the range of about 1 pM to about 100 μM, more usually in the range of about 100 pM to about 10 μM. In specific situations the concentrations may be higher or lower, depending on the nature of the analyte, the affinity of the binding compounds, the efficiency of
10 release of the molecular tags, the sensitivity with which the molecular tags are detected, and the number of analytes to be determined in the assay, as well as other considerations.

In some embodiments, where components of the assay mixture interfere with a chromatographic analysis, the molecular tags may be required to be separated from the assay mixture prior to chromatographic analysis, or certain components of the assay mixture, e.g. binding
15 moieties with unreleased molecular tags, may be required to be excluded from the chromatographic analysis. Depending on the nature of the molecular tags and the components of the assay mixture, one may sequester or adsorb or exclude such binding moieties by using guard column, and the like. Alternatively, one may have a capture ligand attached to binding compounds for the purpose of removing such interfering components in the mixture.
20   An additional degree of flexibility can be conferred on an assay by the stage at which the molecular tags are labeled. A molecular tag may contain a functionality allowing it to bind to a label after reaction with the sample is complete. In this embodiment, a molecular tag comprising a functionality for binding to a detectable label is combined with a sample. After a binding reaction takes place and molecular tags are released, additional reagents are combined in a sample vessel
25 with the products of the first reaction, which react with the released molecular tags to add a detectable label.

For quantitation, one may choose to use controls, which provide a signal in relation to the amount of the target that is present or is introduced. A control to allow conversion of relative fluorescent signals into absolute quantities is accomplished by addition of a known quantity of a
30 fluorophore to each sample before separation of the molecular tags. Any fluorophore that does not interfere with detection of the molecular tag signals can be used for normalizing the fluorescent signal. Such standards preferably have separation properties that are different from those of any of the molecular tags in the sample, and could have the same or a different emission wavelength. Exemplary fluorescent molecules for standards include ROX, FAM, and fluorescein and
35 derivatives thereof.

36

GCPR Pathway Assays

G-protein coupled receptors (GPCRs) represent one of the most important families of drug targets. G protein-mediated signaling systems have been identified in many divergent organisms, such as mammals and yeast. GPCRs respond to, among other extracellular signals,

5   neurotransmitters, hormones, odorants and light. GPCRs are thought to represent a large superfamily of proteins that are characterized by the seven distinct hydrophobic regions, each about 20-30 amino acids in length, that forms the transmembrane domain. The amino acid sequence is not conserved across the entire superfamily, but each phylogenetically related subfamily contains a number of highly conserved amino acid motifs that can be used to identify and classify new

10   members. Individual GPCRs activate particular signal transduction pathways, although at least ten different signal transduction pathways are known to be activated via GPCRs. For example, the beta 2-adrenergic receptor (βAR) is a prototype mammalian GPCR. In response to agonist binding, βAR receptors activate a G protein (G_s) which in turn stimulates adenylate cyclase and cyclic adenosine monophosphate production in the cell.

15   It has been postulated that members of the GPCR superfamily desensitize via a common mechanism involving G protein-coupled receptor kinase (GRK) phosphorylation followed by arrestin binding. The protein β-arrestin regulates GPCR signal transduction by binding agonist-activated receptors that have been phosphorylated by G protein receptor kinases. The β-arrestin protein remains bound to the GPCR during receptor internalization. The interaction between a

20   GPCR and α-arrestin can be measured using several methods. In one example, the β-arrestin protein is fused to green fluorescent protein to create a protein fusion (Barak et al. (1997) *J. Biol. Chem.* 272(44):27497-500). The agonist-dependent binding of β-arrestin to a GPCR can be visualized by fluorescence microscopy. Microscopy can also be used to visualize the subsequent trafficking of the GPCR β-arrestin complex to clathrin coated pits. Other methods for measuring

25   binding of β-arrestin to a GPCR in live cells include techniques such as FRET (fluorescence resonance energy transfer), BRET (bioluminescent energy transfer) or enzyme complementation (Rossi et al. (1997) *Proc. Natl Acad. Sci. U S A* 94(16):8405-10).

At present, there are nearly 400 GPCRs whose natural ligands and function are known. These known GPCRs, named for their endogenous ligands, have been classified into five major

30   categories: Class-A Rhodopsin-like; Class-B Secretin-like; Class-C Metabotropic glutamate/pheromone; Class-D Fungal pheromone; Class-E cAMP (dictyostelium). Representative members of Class-A are the amine receptors (e.g., muscarinic, nicotinic, adrenergic, adenosine, dopamine, histamine and serotonin), the peptide receptors (e.g., angiotensin, bradykinin, chemokines, endothelin and opioid), the hormone receptors (e.g., follicle stimulating, lutropin and

35   thyrotropin), and the sensory receptors, including rhodopsin (light), olfactory (smell) and gustatory (taste) receptors. Representatives of Class-B include secretin, calcitonin, gastrin and glucagon receptors.

37

Many available therapeutic drugs in use today target GPCRs, as they mediate vital physiological responses, including vasodilation, heart rate, bronchodilation, endocrine secretion, and gut peristalsis (Wilson and Bergsma (2000) Pharm. News 7: 105-114). For example, ligands to β-adrenergic receptors are used in the treatment of anaphylaxis, shock, hypertension,

5   hypotension, asthma and other conditions. Additionally, diseases can be caused by the occurrence of spontaneous activation of GPCRs, where a GPCR cellular response is generated in the absence of a ligand. Drugs that are antagonists of GPCRs decrease this spontaneous activity (a process known as inverse agonism) are important therapeutic agents.

Due to the therapeutic importance of GPCRs, methods for the rapid screening of

10 compounds for GPCR ligand activity are desirable. The present invention provides a method of screening test compounds and test conditions for the ability to modulate (activate or inhibit, enhance or depress) a GPCR pathway, and provides methods of assessing GPCR pathway function, such as the function of an orphan GPCR, in a cell in general. In another aspect of the present method, lipophilic photosensitizers are attached to the cellular membranes. A candidate ligand or a

15 library of candidate ligands can be attached to a molecular tag, after which the ligand is allowed to bind to the receptor. After excitation of the photosensitizer with a light source, the cleavable linker is cleaved releasing the molecular tag. The released molecular tag can be detected in the extracellular fluid, as detailed above, which provides information on the structure of the ligand for the GPCR.

20 In one aspect, the present invention provides methods for screening modulators of GPCR activity comprising: a) providing a cell expressing a known or unknown GPCR, wherein the cell is labeled with a lipophilic photosensitizer, b) exposing the cell to a test compound conjugated by a cleavable linkage to a molecular tag; c) illuminating the photosensitizer to generate singlet oxygen that cleaves the molecular tag, d) detecting the signal from the released molecular tag, and (d)

25 comparing the signal produced in the presence of the test compound with the signal produced in the absence, wherein changes in the signal indicates that the compound is a modulator of a GPCR.

The present invention thus provides a convenient method of identifying modulators for an orphan GPCR. Orphan GPCRs are novel receptors typically identified by sequence comparison-based methods, but whose cognate ligands are not known. It is estimated that from 400 to as many as

30 5000 orphan GPCRs may be coded for in the human genome, representing a vast potential for developing new drugs.

Preparation of Cells that Express GPCRs

Methods for preparing cells that express GPCRs have been described. See, e.g., U.S.

35 patents 6,051,386, 6,069,296, 6,111,076 and 6,280,934. Generally, complementary DNA encoding GPCRs can be obtained and can be expressed in an appropriate cell host using techniques well

known in the art. Typically, once a full-length GPCR cDNA has been obtained, it can be expressed in a mammalian cell line, yeast cell, amphibian cell or insect cell for functional analysis. Preferably, the cell line is a mammalian cell line that has been characterized for GPCR expression and that optionally contains a wide repertoire of G-proteins to allow functional coupling to

5   downstream effectors. Examples of such cell lines include Chinese Hamster Ovary (CHO) or Human Embryonic Kidney 293 (HEK293) lines. Cells in which the cDNA is expressed can be encoded using the methods disclosed herein, thus allowing the multiplex screening of ligands. The expressed receptor can then be screened in a variety of functional assays to identify an activating ligand as disclosed above.

10

EXAMPLES

The invention is demonstrated further by the following syntheses and illustrative examples. Parts and percentages are by weight unless otherwise indicated. Temperatures are in degrees Centigrade (°C) unless otherwise specified. The following preparations and examples illustrate the

15   invention but are not intended to limit its scope. Unless otherwise indicated, peptides used in the following examples were prepared by synthesis using an automated synthesizer and were purified by gel electrophoresis or HPLC.

The following abbreviations have the meanings set forth below:

Tris HCl - Tris(hydroxymethyl)aminomethane-HCl (a 10x solution) from BioWhittaker,

20   Walkersville, MD

TLC – thin layer chromatography

BSA - bovine serum albumin, *e.g.* available from Sigma Chemical Company (St. Louis, MO), or like reagent supplier.

EDTA - ethylene diamine tetra-acetate from Sigma Chemical Company

25   FAM - carboxyfluorescein

EMCS - N-ϵ-maleimidocaproyloxy-succinimide ester

EDC - 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

NHS - *N*-hydroxysuccinimide

DCC - 1,3-dicylcohexylcarbodiimide

30   DMF – dimethylformamide

Fmoc – N-(9-fluorenylmethoxycarbonyl)-

Example 1

Conjugation and Release of a Molecular Tag

35   Figure 7A-B summarize the methodology for conjugation of molecular tag precursor to an antibody or other binding compound with a free amino group, and the reaction of the resulting conjugate with singlet oxygen to produce a sulfinic acid moiety as the released molecular tag.

Figure 8 A-J shows several molecular tag reagents, most of which utilize 5- or 6-carboxyfluorescein (FAM) as starting material.

Example 2

Preparation of Pro2, Pro4, and Pro6 through Pro13

The scheme outlined in Figure 9A shows a five-step procedure for the preparation of the carboxyfluorescein-derived molecular tag precursors, namely, Pro2, Pro4, Pro6, Pro7, Pro8, Pro9, Pro10, Pro11, Pro12, and Pro13. The first step involves the reaction of a 5- or 6-FAM with N-hydroxysuccinimide (NHS) and 1,3-dicylcohexylcarbodiimide (DCC) in DMF to give the corresponding ester, which was then treated with a variety of diamines to yield the desired amide, compound 1. Treatment of compound 1 with N-succinimidyl iodoacetate provided the expected iodoacetamide derivative, which was not isolated but was further reacted with 3-mercaptopropionic acid in the presence of triethylamine. Finally, the resulting β-thioacid (compound 2) was converted, as described above, to its NHS ester. The various e-tag moieties were synthesized starting with 5- or 6-FAM, and one of various diamines. The diamine is given $H_2N \wedge X \wedge NH_2$ in the first reaction of Figure 9A. The regioisomer of FAM and the chemical entity of "X" within the diamine are indicated in the table below for each of the molecular tag precursors synthesized. Clearly, the diamine, X, can have a wide range of additional forms, as described above in the discussion of the mobility modifier moiety.

Precursor	FAM	X
Pro2	5-FAM	$C(CH_3)_2$
Pro4	5-FAM	no carbon
Pro6	5-FAM	$(CH_2)_8$
Pro7	5-FAM	$CH_2OCH_2CH_2OCH_2$
Pro8	5-FAM	$CH_2CH_2OCH_2CH_2OCH_2CH_2OCH_2CH_2$
Pro9	5-FAM	1,4-phenyl
Pro10	6-FAM	$C(CH_3)_2$
Pro11	6-FAM	no carbon
Pro12	6-FAM	$CH_2OCH_2CH_2OCH_2$
Pro13	6-FAM	$CH_2CH_2OCH_2CH_2OCH_2CH_2OCH_2CH_2$

Synthesis of compound 1

To a stirred solution of 5- or 6-carboxyfluorescein (0.5 mmol) in dry DMF (5 mL) were added N-hydroxysuccinimide (1.1 equiv.) and 1,3-dicylcohexylcarbodiimide (1.1 equiv.). After about 10 minutes, a white solid (dicyclohexylurea) started forming. The reaction mixture was

stirred under nitrogen at room temperature overnight. TLC (9:1 CH_2Cl_2-MeOH) indicated complete disappearance of the starting material.

The supernatant from the above mixture was added dropwise to a stirred solution of diamine (2-5 equiv.) in DMF (10 mL). As evident from TLC (40:9:1 CH_2Cl_2-MeOH-H_2O), the
5 reaction was complete instantaneously. The solvent was removed under reduced pressure. Flash chromatography of the resulting residue on Iatrobeads silica provided the desired amine (compound 1) in 58-89% yield. The ^{1}H NMR (300 MHz, DMSO-d_6) of compound 1 was in agreement with the assigned structure.

Synthesis of compound 2

10 To the amine (compound 1) (0.3 mmol) were sequentially added dry DMF (10 mL) and N-succinimidyl iodoacetate (1.1 equiv.). The resulting mixture was stirred at room temperature until a clear solution was obtained. TLC (40:9:1 CH_2Cl_2-MeOH-H_2O) revealed completion of the reaction.

The above reaction solution was then treated with triethylamine (1.2 equiv.) and 3-
15 mercaptopropionic acid (3.2 equiv.). The mixture was stirred at room temperature overnight. Removal of the solvent under reduced pressure followed by flash chromatography afforded the β-thioacid (compound 2) in 62-91% yield. The structure of compound 2 was assigned on the basis of its ^{1}NMR (300 MHz, DMSO-d_6).

20 ### Synthesis of Pro2, Pro4, and Pro6 through Pro13

To a stirred solution of the β-thioacid (compound 2) (0.05 mmol) in dry DMF (2 mL) were added N-hydroxysuccinimide (1.5 equiv.) and 1,3-dicylcohexylcarbodiimide (1.5 equiv.). The mixture was stirred at room temperature under nitrogen for 24-48 h (until all of the starting material had reacted). The reaction mixture was concentrated under reduced pressure and then purified by
25 flash chromatography to give the target molecule in 41-92% yield.

Preparation of Pro1

The compounds of this reaction are shown in Figure 9B. To a stirred solution of 5-iodoacetamidofluorescein (compound 4) (24 mg, 0.047 mmol) in dry DMF (2 mL) were added
30 triethylamine (8 µL, 0.057 mmol) and 3-mercaptopropionic acid (5µL, 0.057 mmol). The resulting solution was stirred at room temperature for 1.5 h. TLC (40:9:1 CH_2Cl_2-MeOH-H_2O) indicated completion of the reaction. Subsequently, N-hydroxysuccinimide (9 mg, 0.078 mmol) and 1,3-dicylcohexylcarbodiimide (18 mg, 0.087 mmol) were added. The reaction mixture was stirred at room temperature under nitrogen for 19 h at which time TLC showed complete disappearance of
35 the starting material. Removal of the solvent under reduced pressure and subsequent flash chromatography using 25:1 and 15:1 CH_2Cl_2-MeOH as eluant afforded Pro1 (23 mg, 83%).

41

Preparation of Pro3

The compounds of this reaction are shown in Figure 9C. To a stirred solution of 6-iodoacetamidofluorescein (compound 5) (26 mg, 0.050 mmol) in dry DMF (2 mL) were added triethylamine (8 μL, 0.057 mmol) and 3-mercaptopropionic acid (5μL, 0.057 mmol). The resulting
5 solution was stirred at room temperature for 1.5 h. TLC (40:9:1 CH_2Cl_2-MeOH-H_2O) indicated completion of the reaction. Subsequently, N-hydroxysuccinimide (11 mg, 0.096 mmol) and 1,3-dicylcohexylcarbodiimide (18 mg, 0.087 mmol) were added. The reaction mixture was stirred at room temperature under nitrogen for 19 h at which time TLC showed complete disappearance of the starting material. Removal of the solvent under reduced pressure and subsequent flash
10 chromatography using 30:1 and 20:1 CH_2Cl_2-MeOH as eluant provided **Pro3** (18 mg, 61%).

Preparation of Pro5

The compounds of this reaction are shown in Figure 9D.

15 ## Synthesis of compound 7

To a stirred solution of 5-(bromomethyl)fluorescein (compound 6) (40 mg, 0.095 mmol) in dry DMF (5 mL) were added triethylamine (15 μL, 0.108 mmol) and 3-mercaptopropionic acid (10μL, 0.115 mmol). The resulting solution was stirred at room temperature for 2 days. TLC (40:9:1 CH_2Cl_2-MeOH-H_2O) indicated completion of the reaction. The reaction solution was
20 evaporated under reduced pressure. Finally, flash chromatography employing 30:1 and 25:1 CH_2Cl_2-MeOH as eluant provided the β-thioacid (compound 7) (28 mg, 66%).

Synthesis of Pro5

To a solution of the acid (compound 7) (27 mg, 0.060 mmol) in dry DMF (2 mL) were
25 added N-hydroxysuccinimide (11 mg, 0.096 mmol) and 1,3-dicylcohexylcarbodiimide (20 mg, 0.097 mmol). The reaction mixture was stirred at room temperature under nitrogen for 2 days at which time TLC (9:1 CH_2Cl_2-MeOH) showed complete disappearance of the starting material. Removal of the solvent under reduced pressure and subsequent flash chromatography with 30:1 CH_2Cl_2-MeOH afforded Pro5 (24 mg, 73%).
30

Preparation of Pro14

The compounds of this reaction are shown in Figure 9E.

Synthesis of compound 9

35 To 5-aminoacetamidofluorescein (compound 8) (49 mg, 0.121 mmol) were sequentially added dry DMF (4 mL) and N-succinimidyl iodoacetate (52 mg, 0.184). A clear solution resulted and TLC (40:9:1 CH_2Cl_2-MeOH-H_2O) indicated complete disappearance of the starting material.

The above reaction solution was then treated with triethylamine (30 µL, 0.215 mmol) and 3-mercaptopropionic acid (30µL, 0.344 mmol). The resulting mixture was stirred for 2 h. Removal of the solvent under reduced pressure followed by flash chromatography using 20:1 and 15:1 CH$_2$Cl$_2$-MeOH as eluant gave the β-thioacid (compound 9) (41 mg, 62%). The structural assignment was made on the basis of ^{1}NMR (300 MHz, DMSO-d_6).

Synthesis of Pro14

To a stirred solution of compound 9 (22 mg, 0.04 mmol) in dry DMF (2 mL) were added N-hydroxysuccinimide (9 mg, 0.078 mmol) and 1,3-dicylcohexylcarbodiimide (16 mg, 0.078 mmol). The resulting solution was stirred at room temperature under nitrogen for about 24 h. The reaction mixture was concentrated under reduced pressure and the residue purified by flash chromatography using 30:1 and 20:1 CH$_2$Cl$_2$-MeOH as eluant to give Pro14 (18 mg, 70%).

Synthesis of Pro15, Pro20, Pro22, and Pro28

The synthesis schemes for producing NHS esters of molecular tags Pro15, Pro20, Pro22, and Pro28 are shown in Figures 16 F-I, respectively. All of the reagent and reaction conditions are conventional in the art and proceed similarly as the reactions described above.

WHAT IS CLAIMED IS:

1.   A method for determining the structure of a ligand for a G protein coupled receptor, the method comprising the steps of:

5   providing a sensitizer-treated membrane comprising the one or more G protein coupled receptors;

providing a plurality of libraries of binding compounds, each library of binding compound having one or more molecular tags, each molecular tag being attached by a cleavable linkage, and the molecular tags of each library of binding compounds being distinguishable from those of every

10   other library of binding compounds by one or more physical and/or optical characteristics;

combining the sensitizer-treated membrane and the plurality of libraries of binding compounds specific such that in the presence of a G protein coupled receptor a complex is capable of being formed between such G protein coupled receptor and members of the libraries of binding compounds;

15   activating the sensitizer-treated membrane so that an active species is produced that cleaves the cleavable linkage of the binding compounds forming such complex so that molecular tags are released; and

separating and identifying the released molecular tags by the one or more physical characteristics to determine the relative binding strength of each library for the G protein coupled

20   receptor.

2.   The method of claim 1 wherein said step of separating includes electrophoretically separating said released molecular tags, and wherein each of said molecular tags has a molecular weight in the range of from 100 to 2500 daltons.

25

3.   The method of claim 2 (a) wherein said sensitizer-treated membrane is a photosensitizer-treated membrane, (b) wherein said active species is singlet oxygen, and (c) wherein each of said molecular tags attached to said binding compounds are selected from a group defined by the formula:

30

$$-L-(M,D)$$
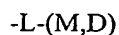
wherein:

L is a cleavable linkage;

35   D is a detection moiety; and

44

M is a bond or a water soluble organic compound consisting of from 1 to 100 atoms, not including hydrogen, that are selected from the group consisting of carbon, oxygen, nitrogen, phosphorus, boron, and sulfur.

4.   The method of claim 3 wherein D is a fluorescent label, a chromogenic label, or an electrochemical label.

5.   The method of claim 4 wherein M is a polymer selected from any one of polyethers, polyesters, polypeptides, oligosaccharides, polyurethanes, polyamides, polysulfonamides, polysulfoxides, polyphosphonates, and block copolymers thereof.

6.   The method of claim 5 wherein D is a fluorescein.

7.   The method of claim 6 wherein said fluorescein is selected from the group consisting of 5- and 6-carboxyfluorescein, 5- and 6-carboxy-4,7-dichlorofluorescein, 2',7'-dimethoxy-5- and 6-carboxy-4,7-dichlorofluorescein, 2',7'-dimethoxy-4',5'-dichloro-5- and 6-carboxyfluorescein, 2',7'-dimethoxy-4',5'-dichloro-5- and 6-carboxy-4,7-dichlorofluorescein, 1',2',7',8'-dibenzo-5- and 6-carboxy-4,7-dichlorofluorescein, 1',2',7',8'-dibenzo-4',5'-dichloro-5- and 6-carboxy-4,7-dichlorofluorescein, 2',7'-dichloro-5- and 6-carboxy-4,7-dichlorofluorescein, and 2',4',5',7'-tetrachloro-5- and 6-carboxy-4,7-dichlorofluorescein.

8.   The method of claim 3 wherein L is selected from the group consisting of olefins, thioethers, selenoethers, thiazoles, oxazoles, and imidazoles.

9.   The method in accordance with claims 1, 2, 3, 4, 5, 6, 7, or 8 wherein said plurality of molecular tags is in the range of from 2 to 50 and wherein each of said plurality of said libraries of binding compounds is a peptide library.

RECEPTOR DEORPHANIZATION USING TAGGED MOLECULAR LIBRARIES

5

Abstract of the Invention

10   The present invention provides a method of screening test compounds and test conditions for the ability to modulate (activate or inhibit, enhance or depress) a GPCR pathway, and provides methods of assessing GPCR pathway function, such as the function of an orphan GPCR, in a cell in general. In another aspect of the present method, lipophilic photosensitizers are attached to the cellular membranes. A candidate ligand or a library of candidate ligands can be attached to a

15   molecular, after which the ligand is allowed to bind to the receptor. After excitation of the photosensitizer with a light source, the cleavable linker is cleaved releasing the molecular tag. The released molecular tag can be detected in the extracellular fluid, as detailed above, which provides information on the structure of the ligand for the GPCR.

Fig. 1A

Fig. 1B

130

mT₁-XX₁-AA₂-AA₃-AA₄- ... -AAₙ

132

mT₂-AA₁-XX₂-AA₃-AA₄- ... -AAₙ

134

mT₃-AA₁-AA₂-XX₃-AA₄- ... -AAₙ

mT₄-AA₁-AA₂-AA₃-XX₄- ... -AAₙ

.
.
.

mTₖ-AA₁-AA₂-AA₃-AA₄- ... -AAₙ
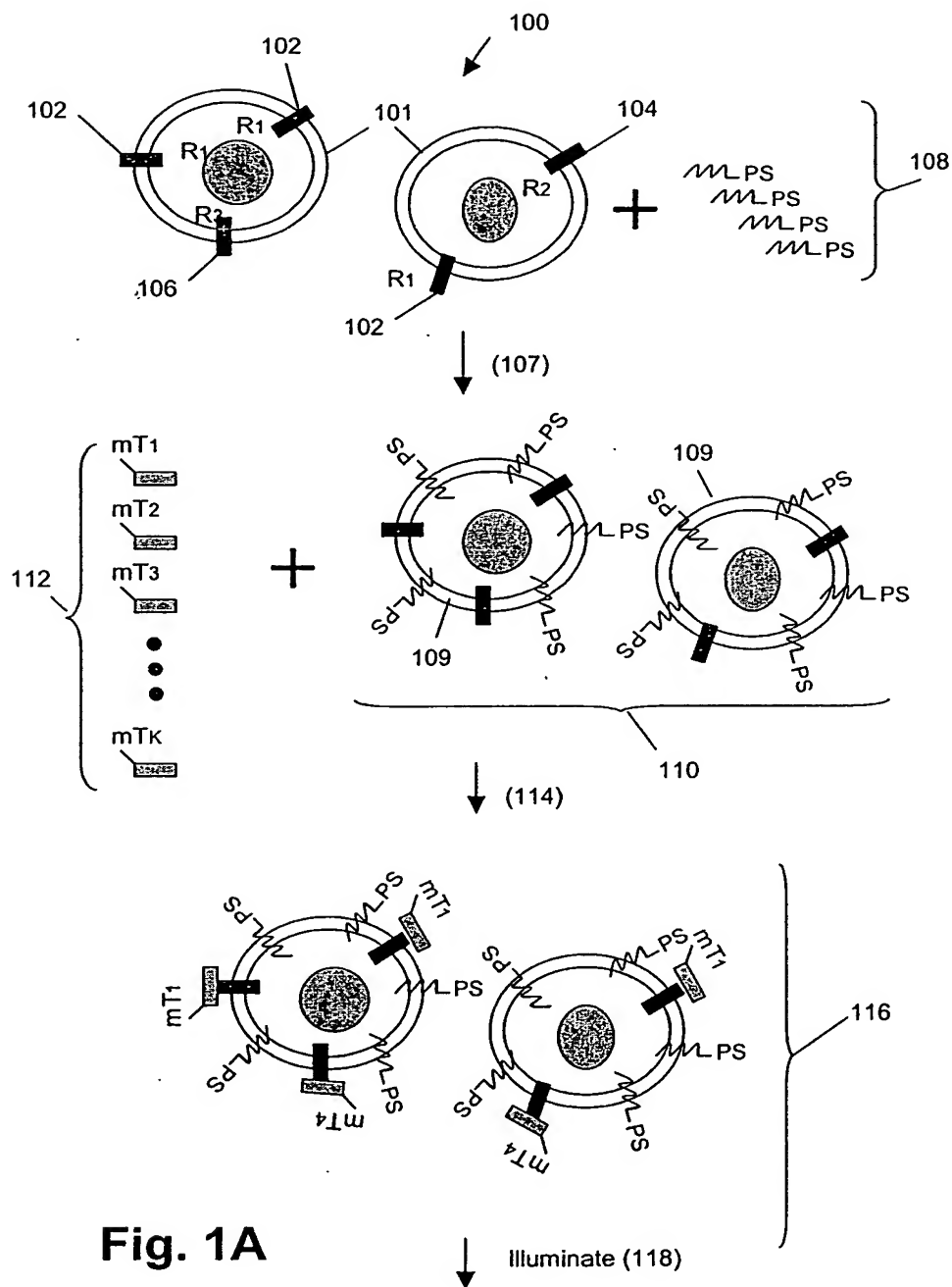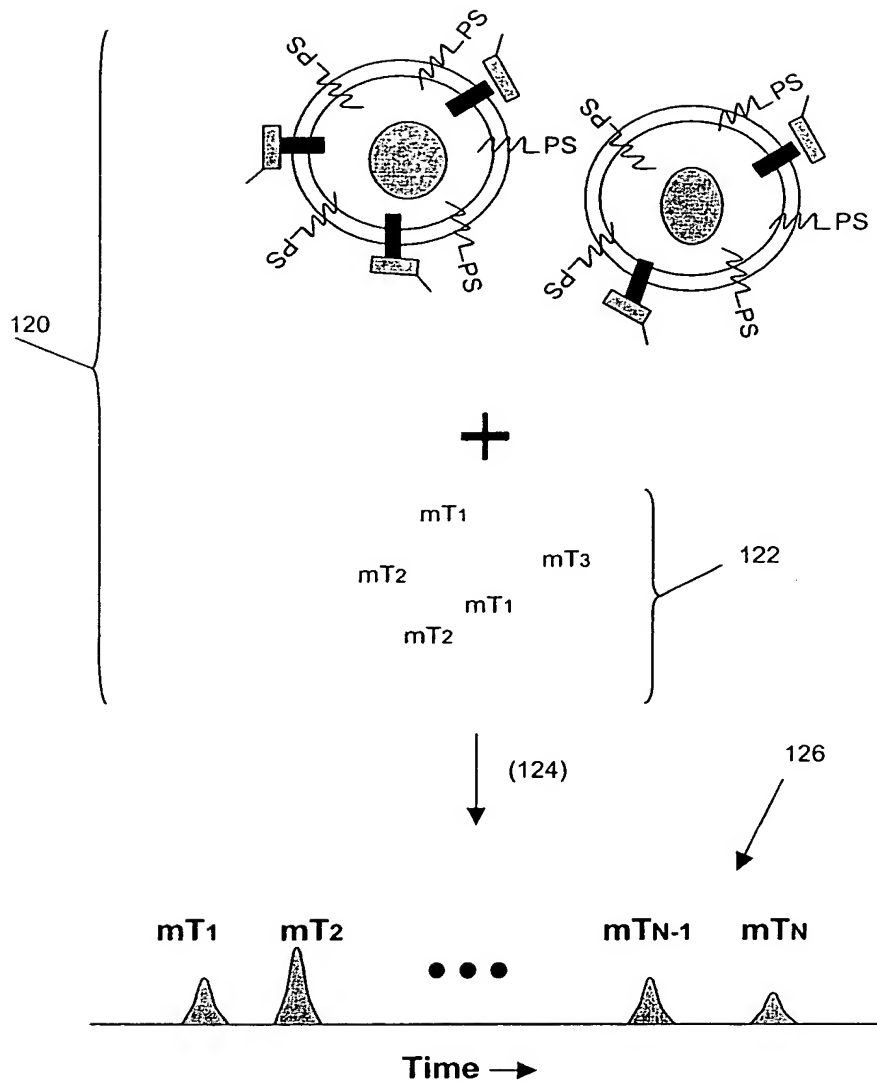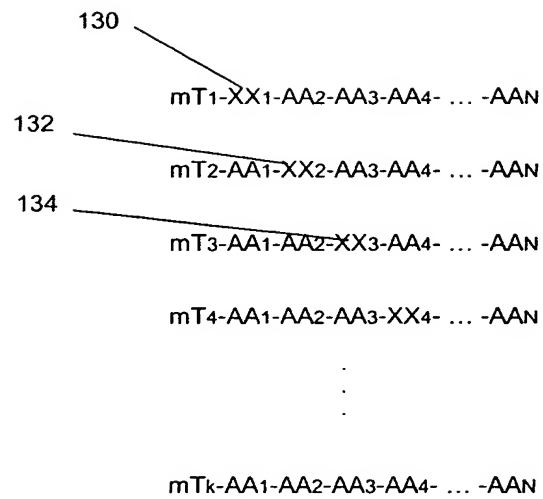
Fig. 1C
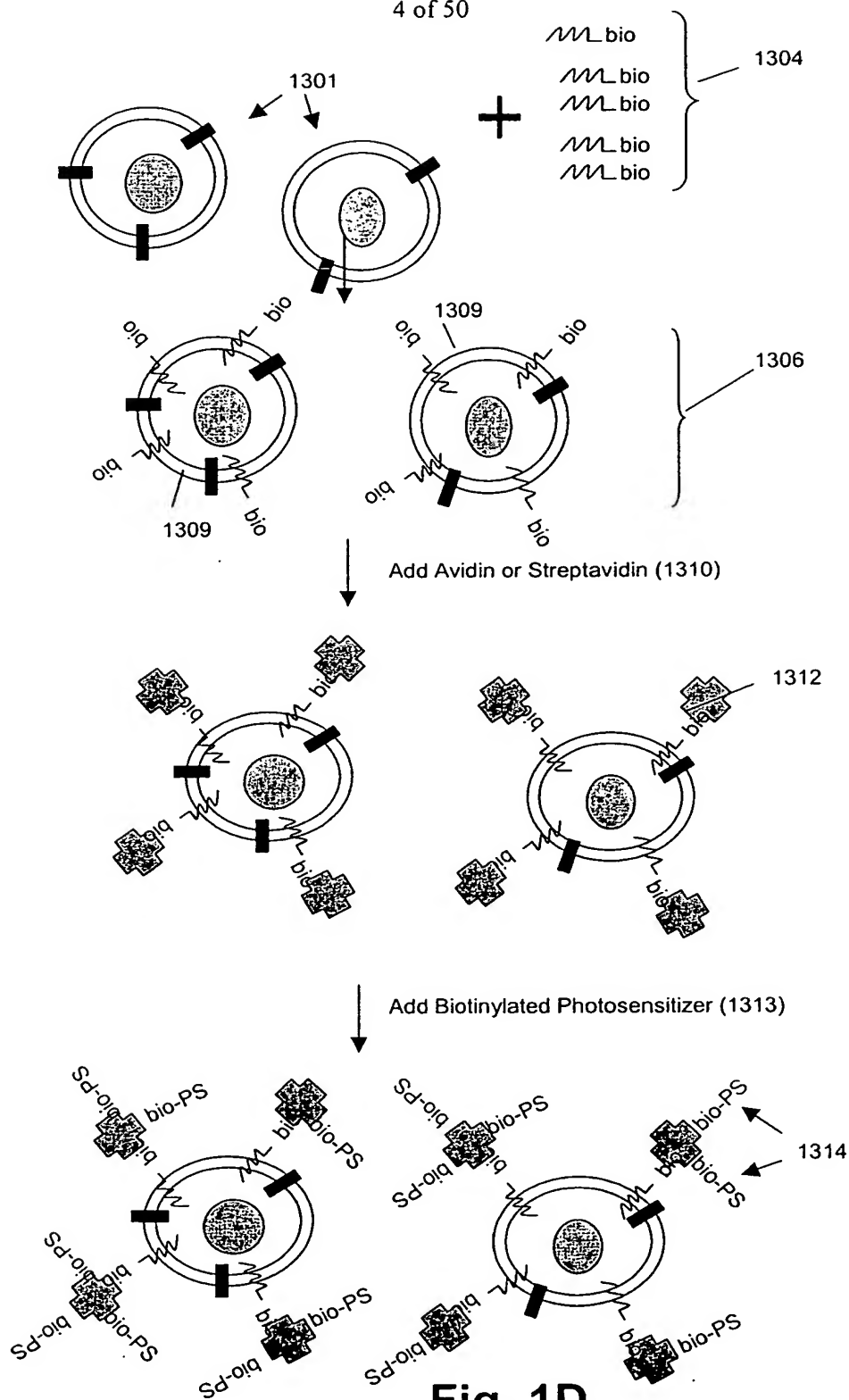
Fig. 1D
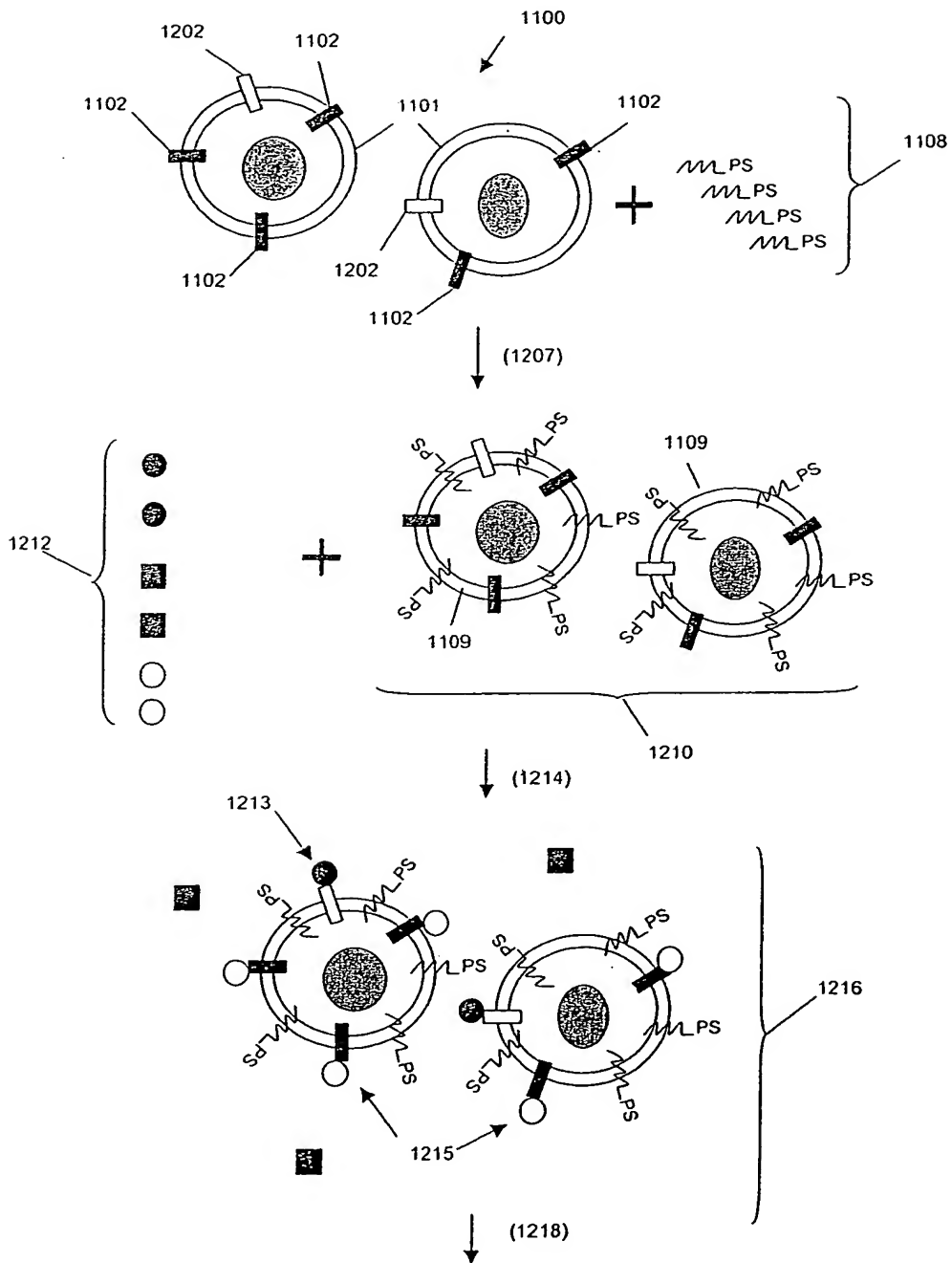
Fig. 1E
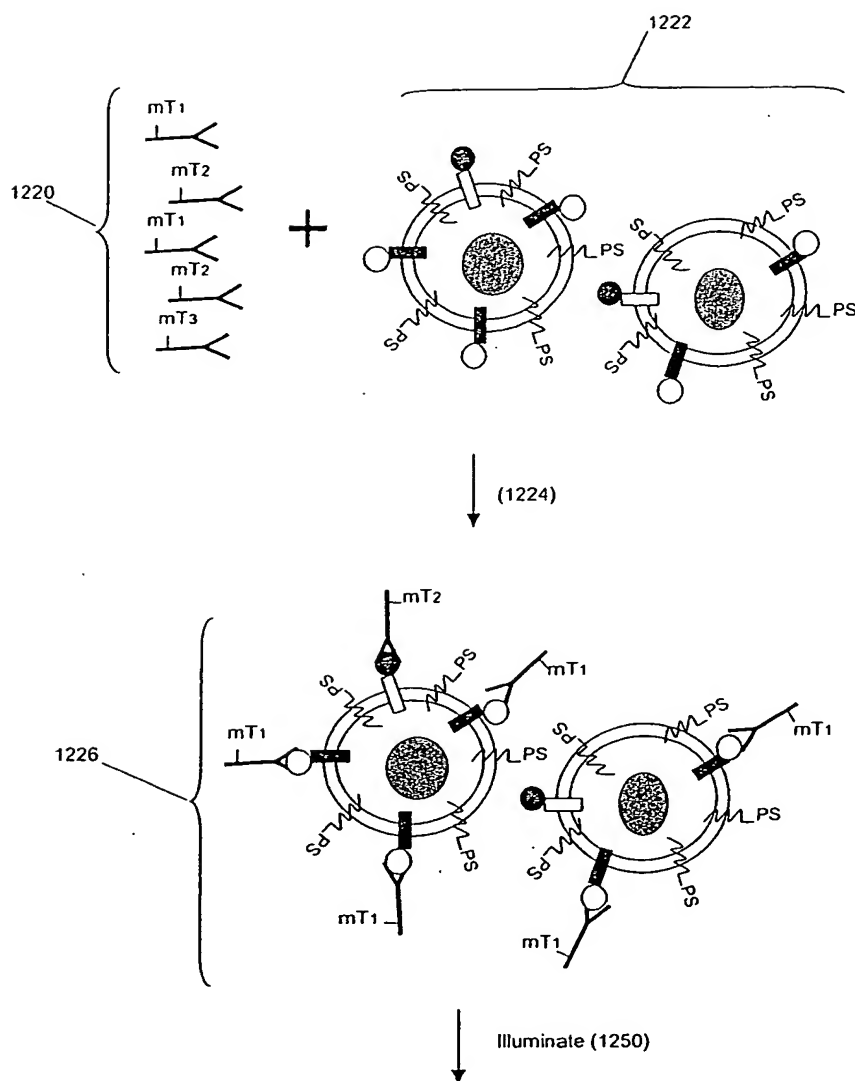
Fig. 1F
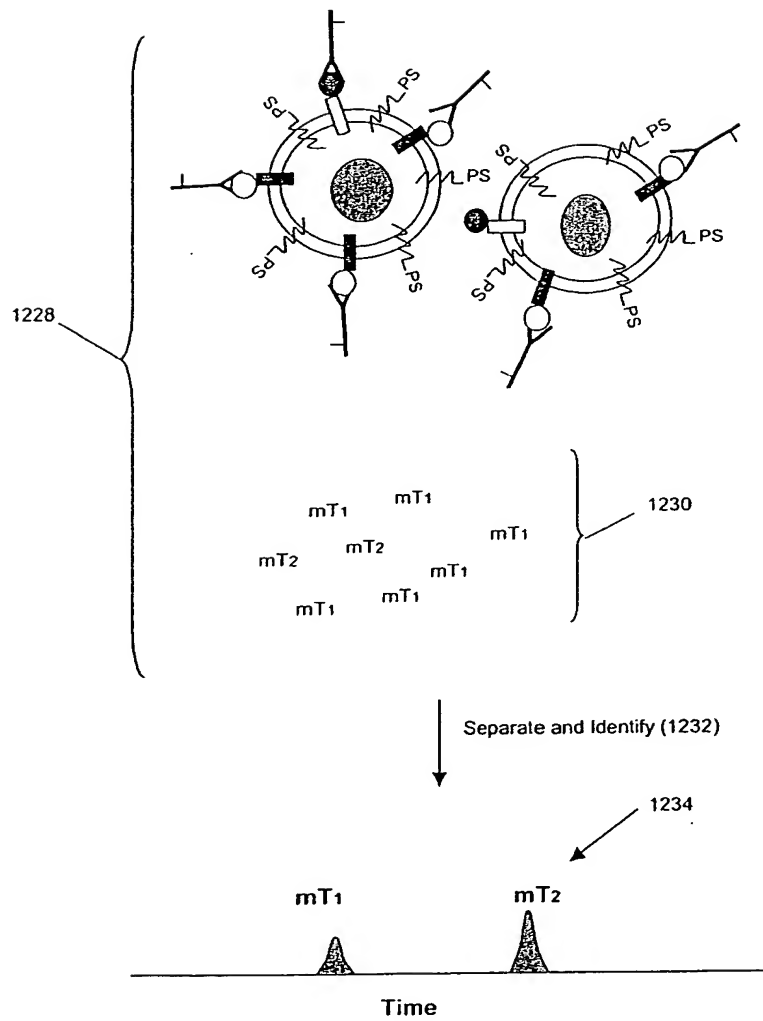
Fig. 1G
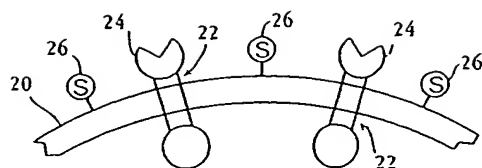
Fig. 1H

Fig. 1I

Fig. 1J

Fig. 1K

Fig. 1L

Fig. 1M

Fig. 1N

Fig. 1O
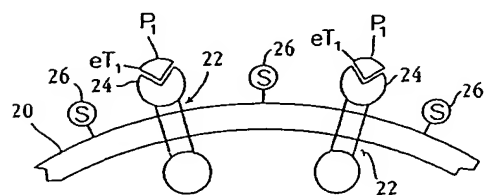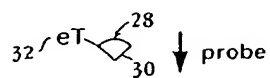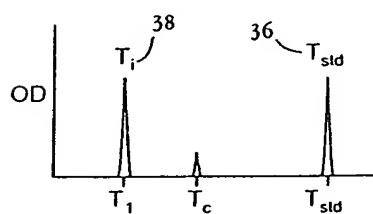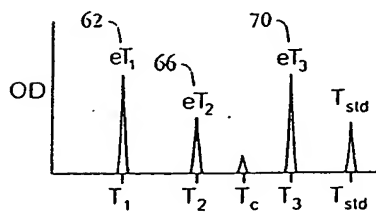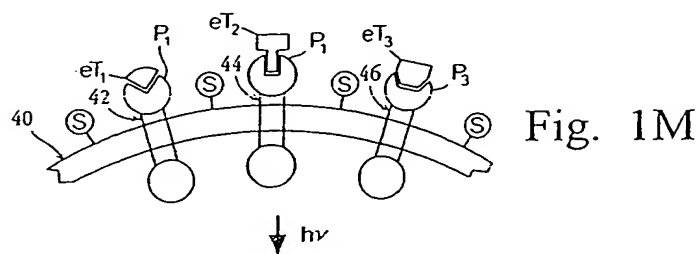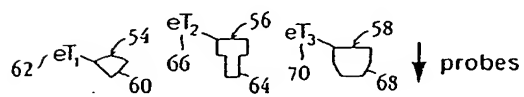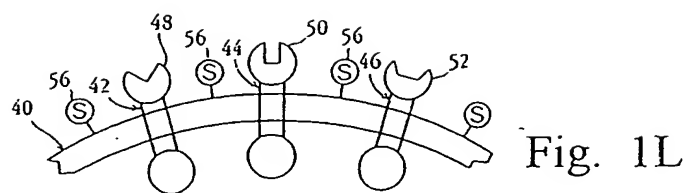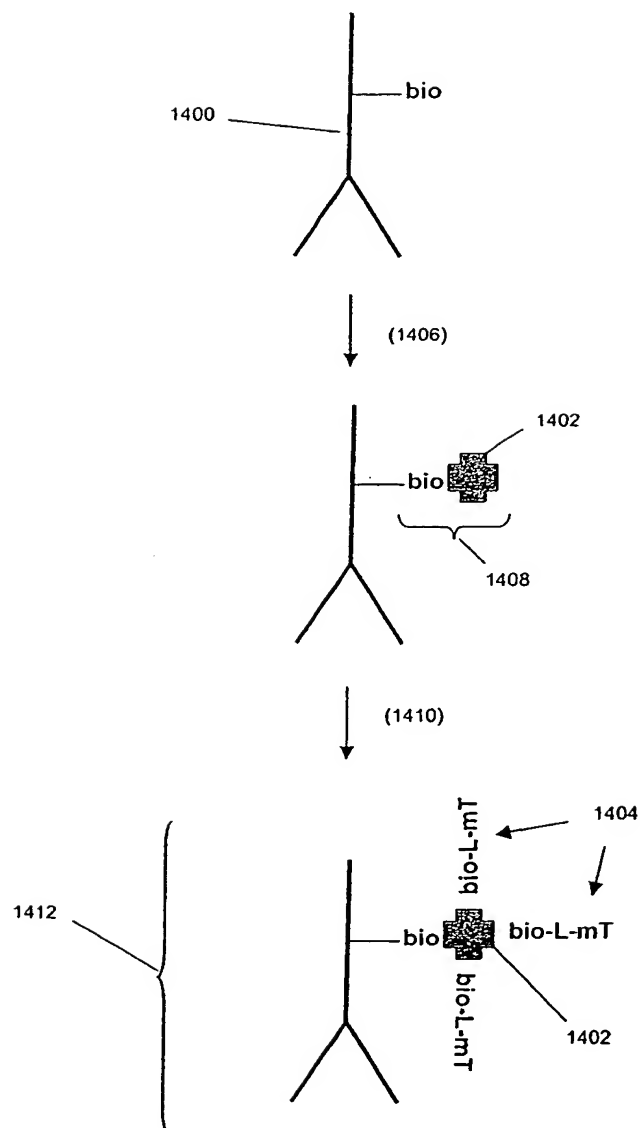
Fig. 1P
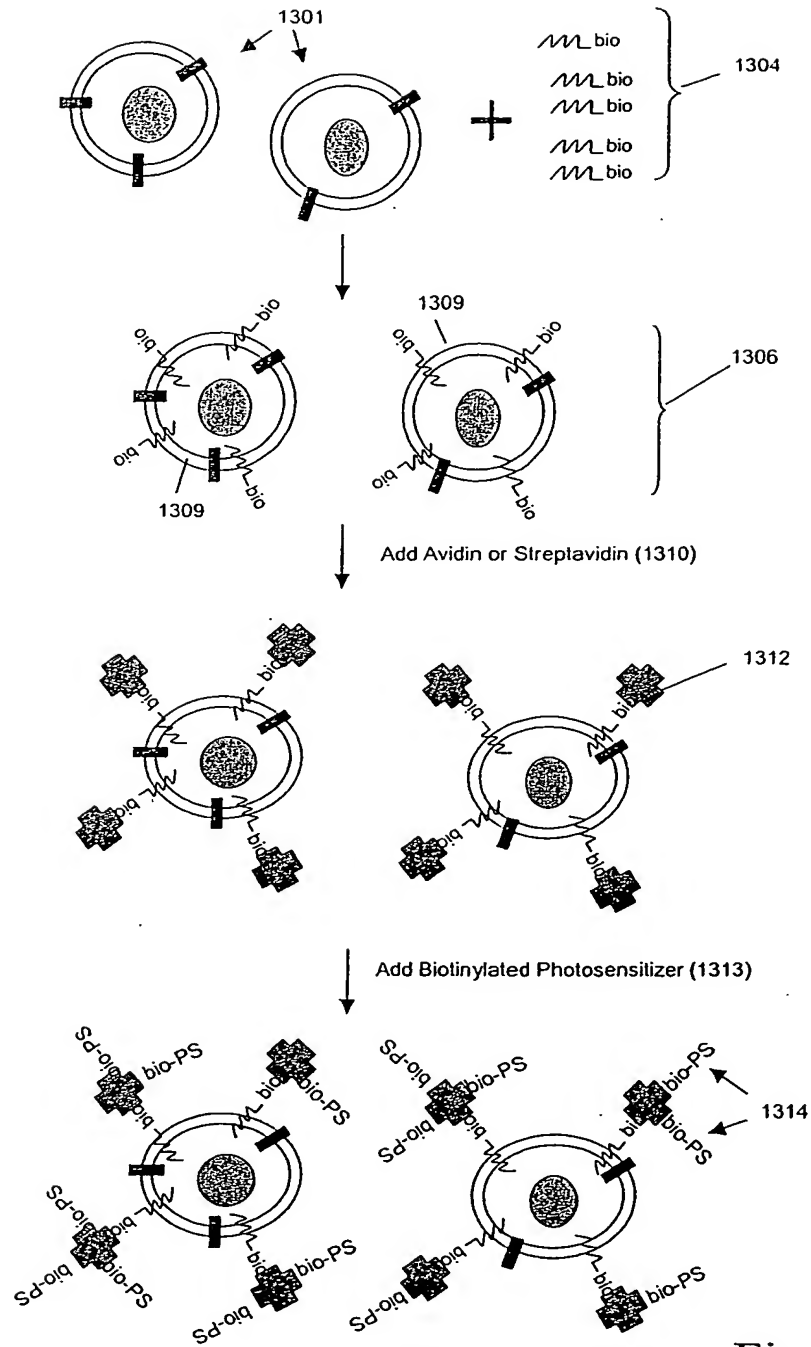
Fig. 1Q
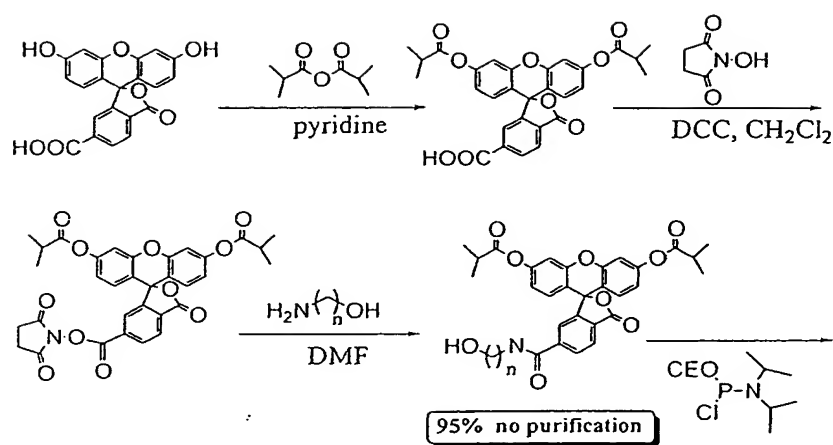
Fig. 2
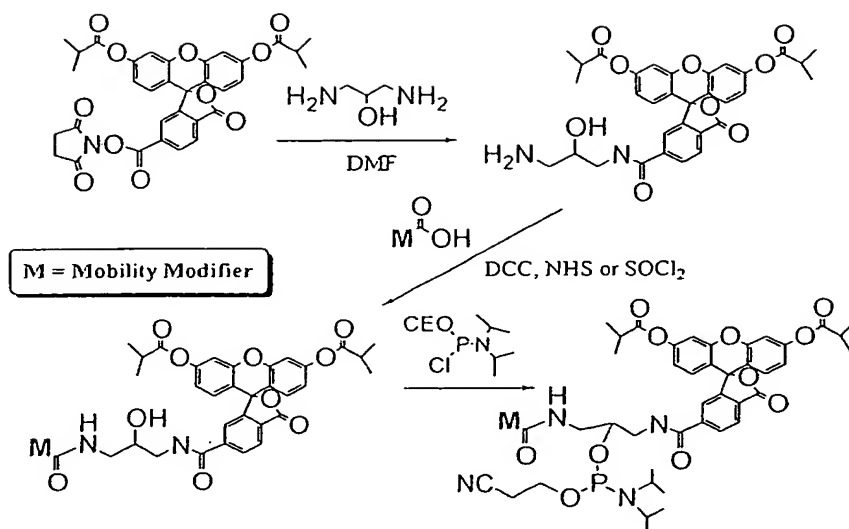
Fig. 3
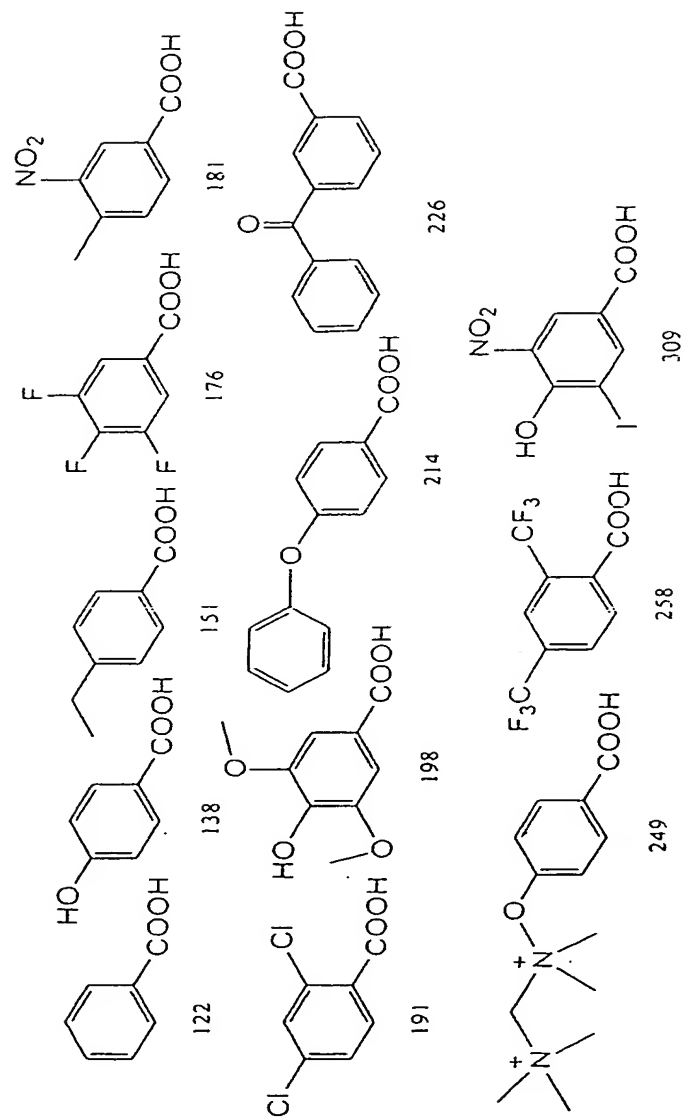
FIG. 4
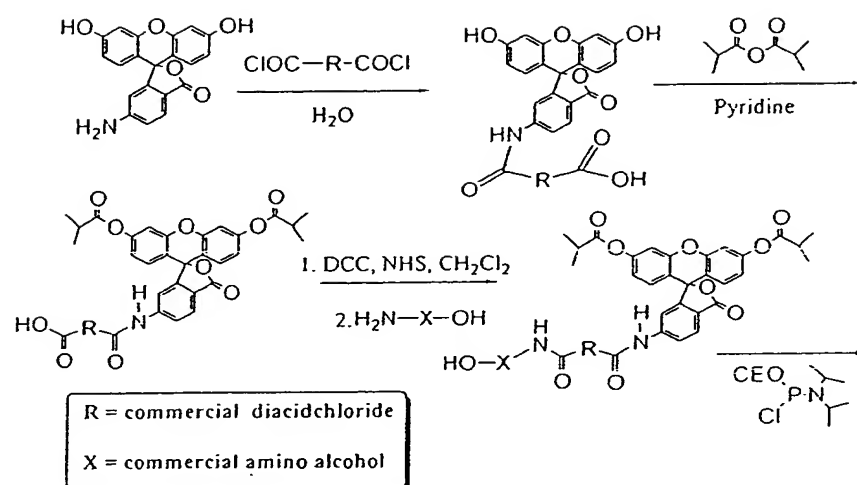
FIG. 5
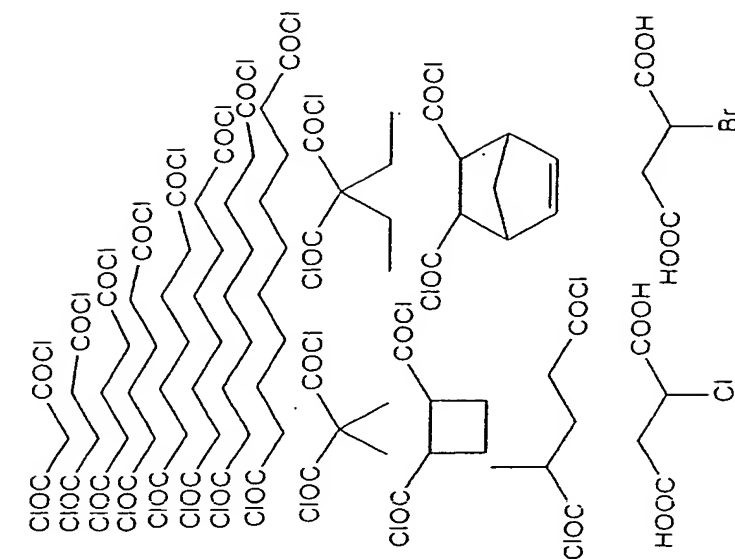
FIG. 6
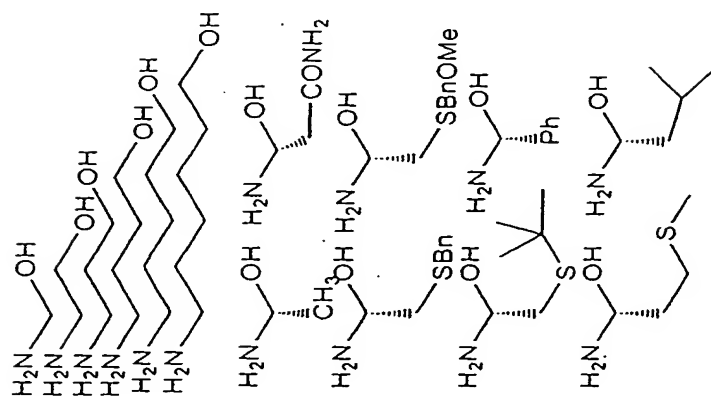
Thiazole cleavable linkage
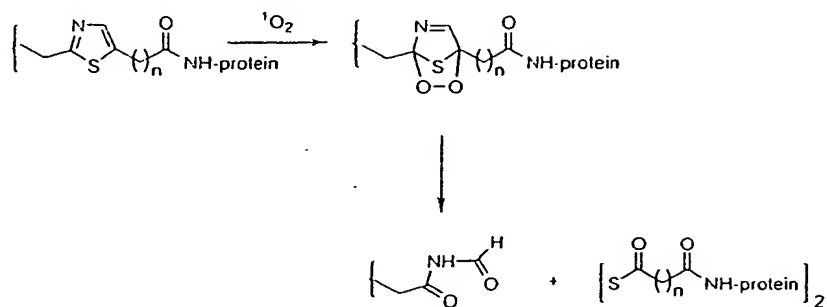
FIG. 7A

Oxazole cleavable linkage
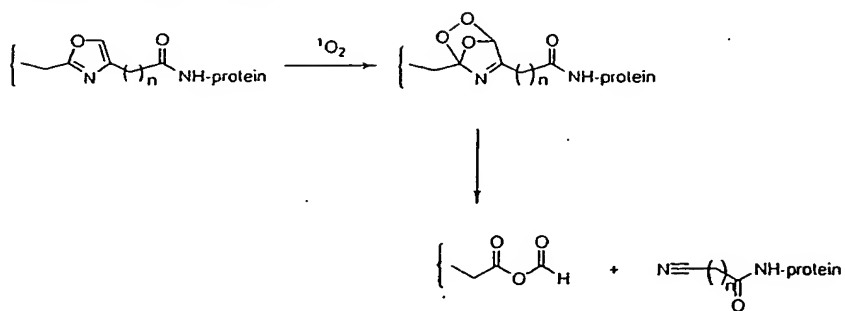
FIG. 7B

Olefin cleavable linkage
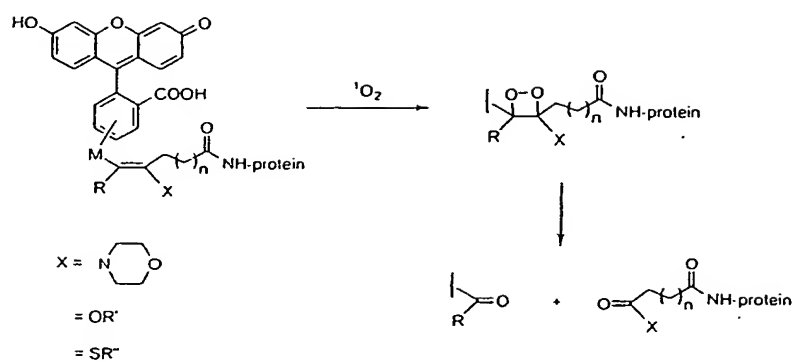
FIG. 7C

Thioether cleavable linkage

FIG. 7D

FIG. 7E

FIG. 7F

M = mobility modifier

released e-tag reporter
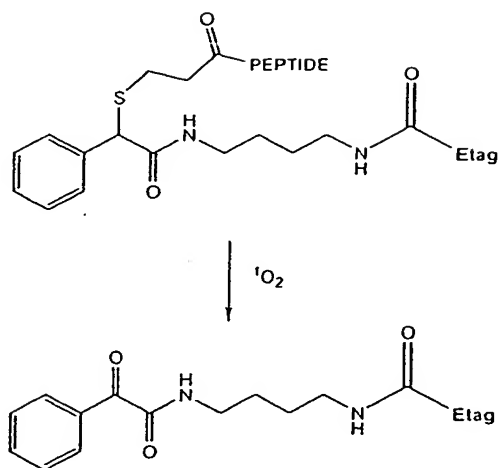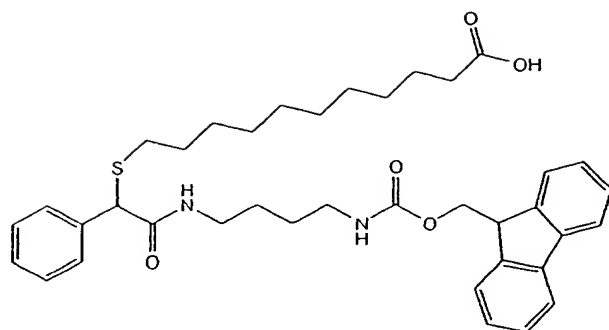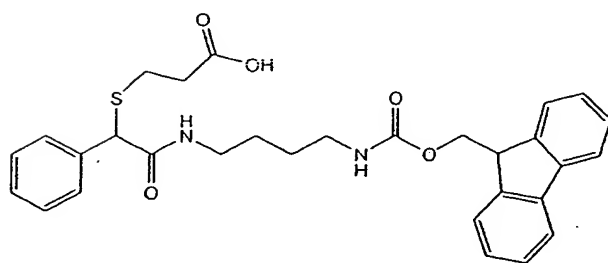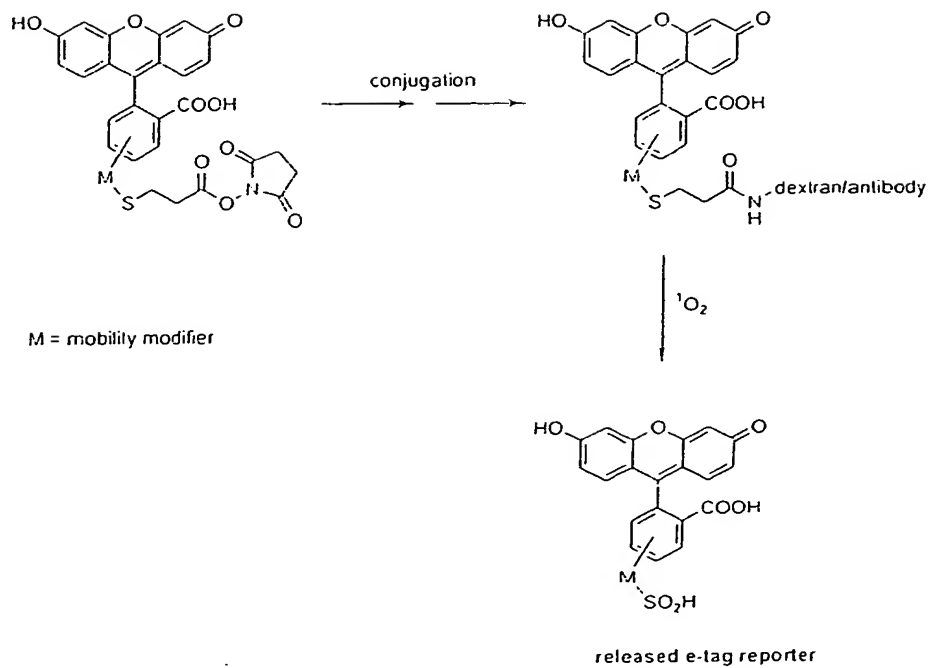
FIG. 8A

Synthesis of FAM-derived eTags

5/6-FAM

1. NHS, DCC, DMF
2. H₂N‿‿X‿NH₂
   DMF
3. bromoacetic anhydride
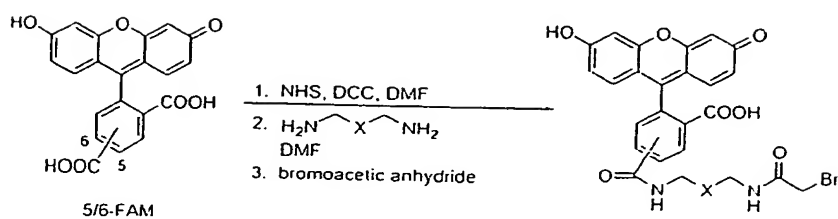
1. 3-mercaptopropionic acid
   Et₃N, DMF
2. NHS, DCC
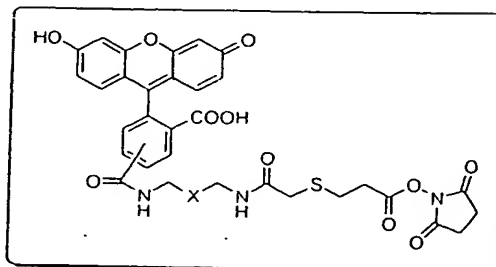
FIG. 8B
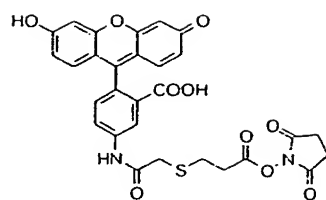
Pro1-NHS
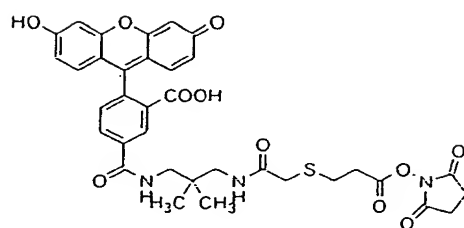
Pro2-NHS
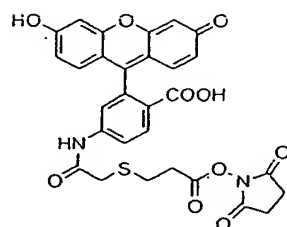
Pro3-NHS
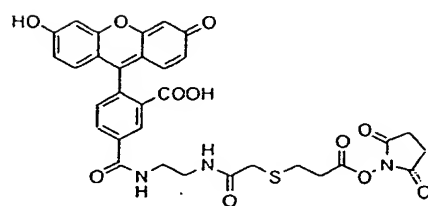
Pro4-NHS

FIG. 9A
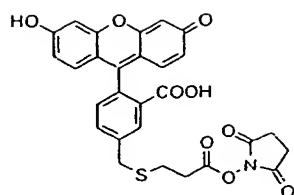
Pro5-NHS
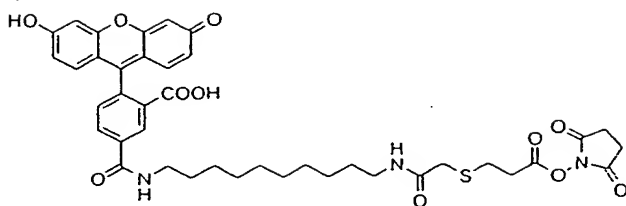
Pro6-NHS
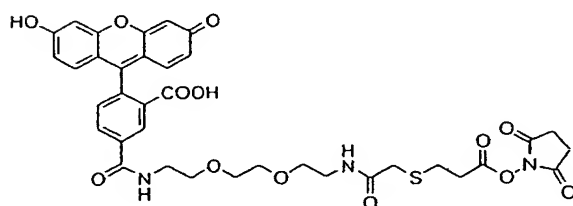
Pro7-NHS
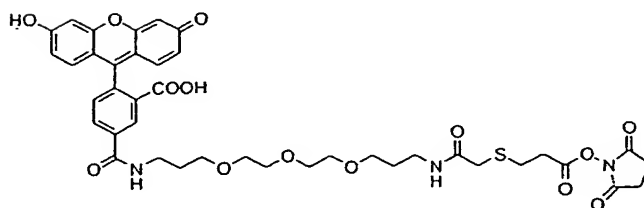
Pro8-NHS

FIG. 9B
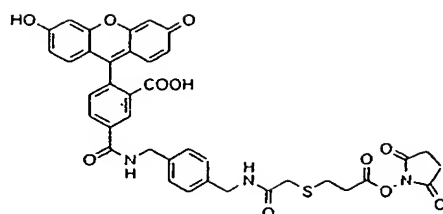
Pro9-NHS
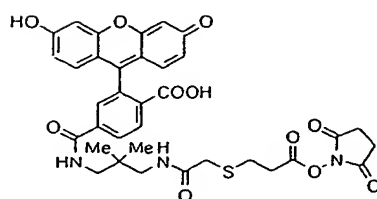
Pro10-NHS
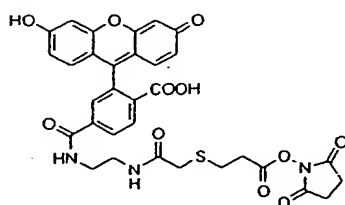
Pro11-NHS

Pro12-NHS

FIG. 9C·

Pro13-NHS

Pro14-NHS

Pro15-NHS

Pro16-NHS
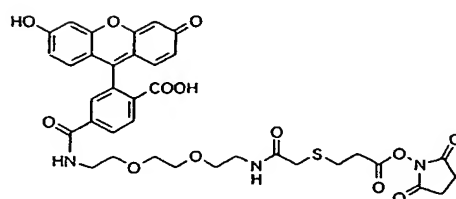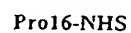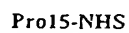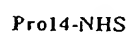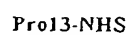
FIG. 9D
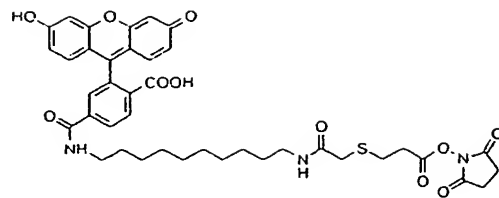
Pro17-NHS
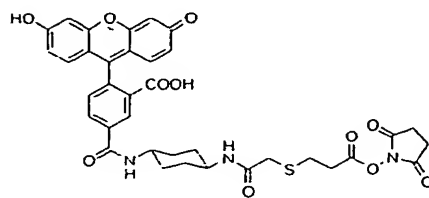
Pro18-NHS
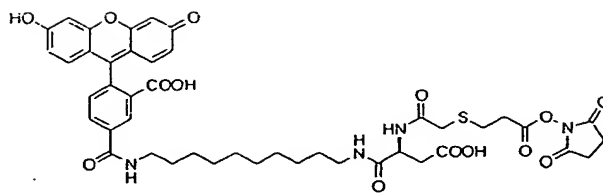
Pro19-NHS

Pro20-NHS

FIG. 9E

Pro21-NHS

Pro22-NHS

Pro23-biotin

Pro24-biotin
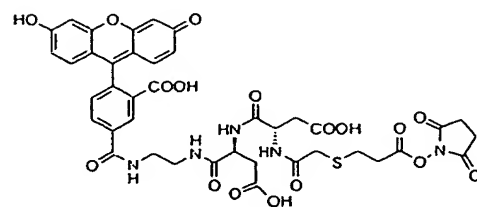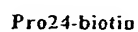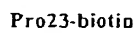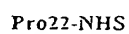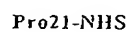
FIG. 9F
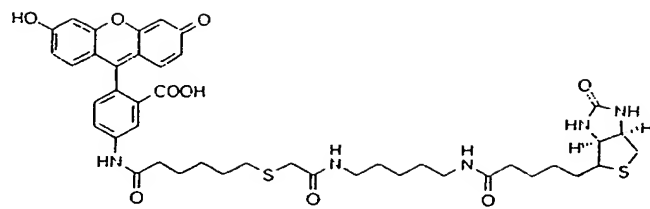
Pro25-biotin
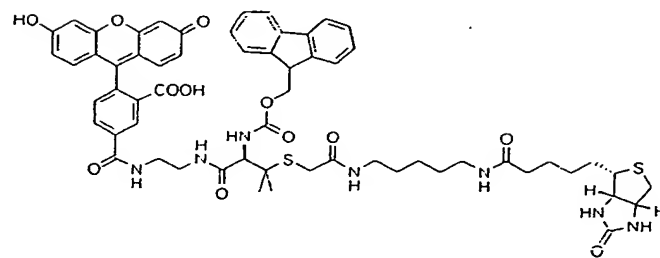
Pro26-biotin
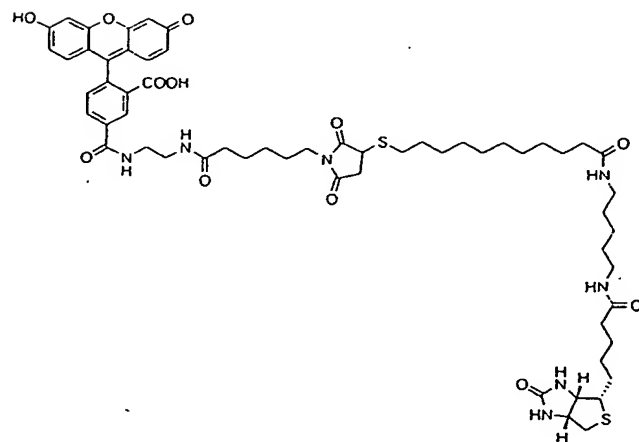
Pro27-biotin

FIG. 9G
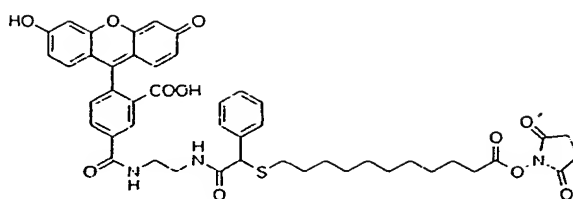
Pro28-NHS
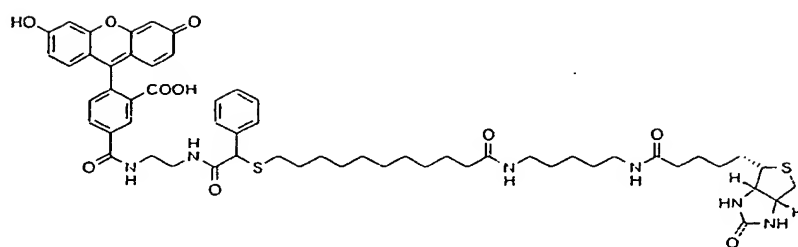
Pro28-biotin
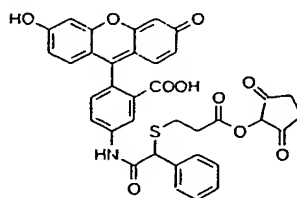
Pro29-NHS
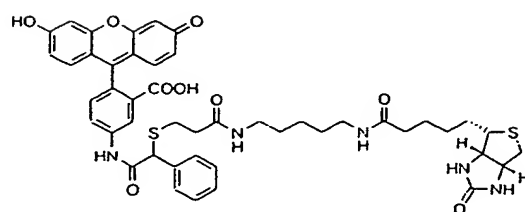
Pro29-biotin

FIG. 9H
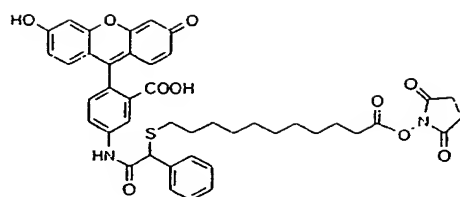
Pro30-NHS
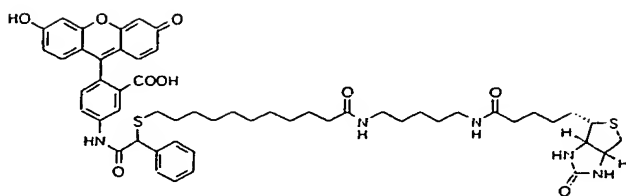
Pro30-biotin
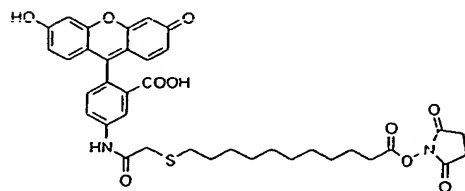
Pro31-NHS
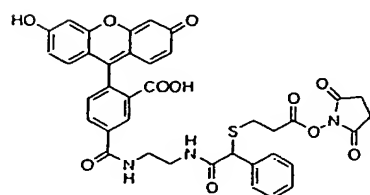
Pro32-NHS

FIG. 9i
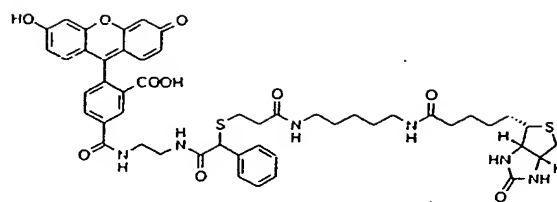
Pro32-biotin
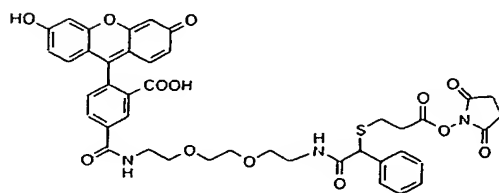
Pro33-NHS
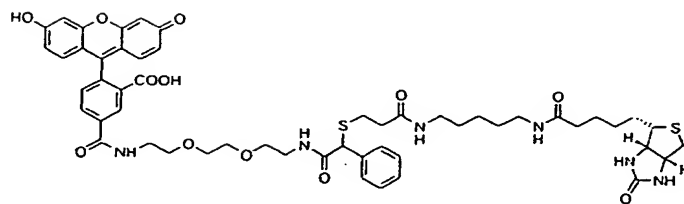
Pro33-biotin

FIG. 9J
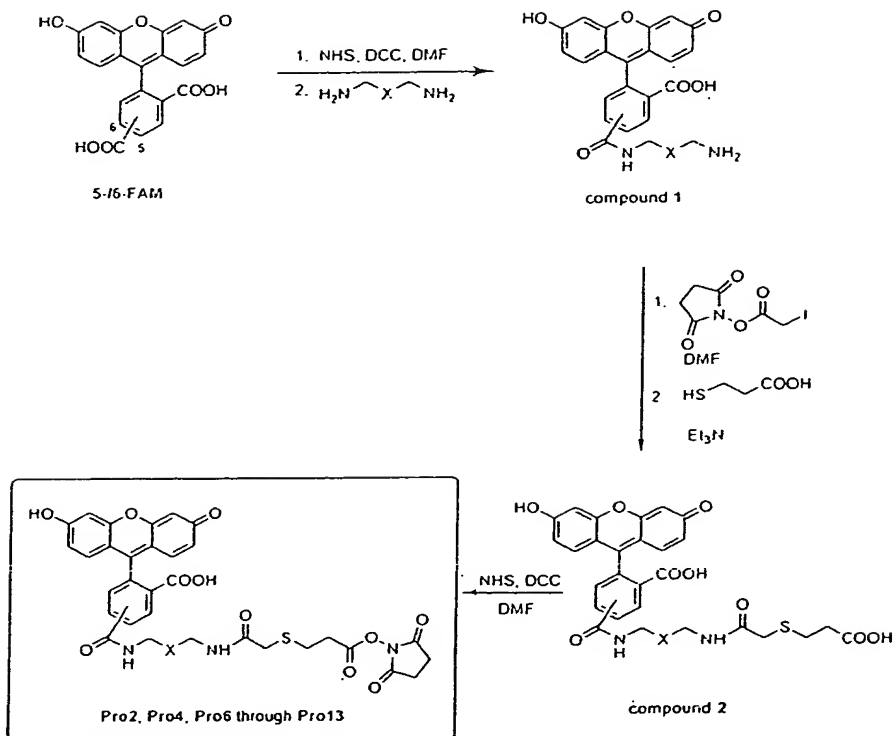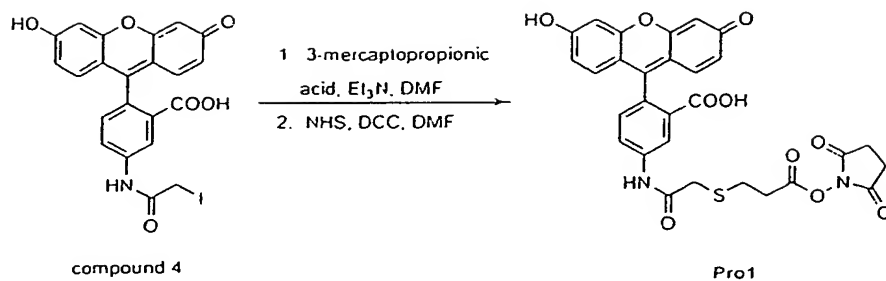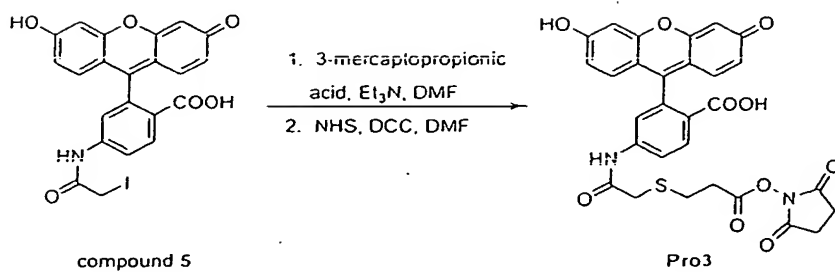
FIG. 10A

FIG. 10B

FIG. 10C
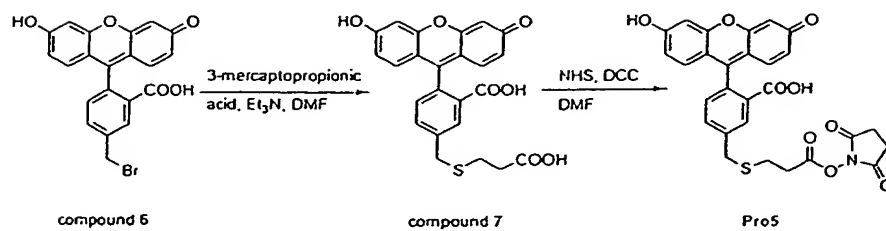
FIG. 10D
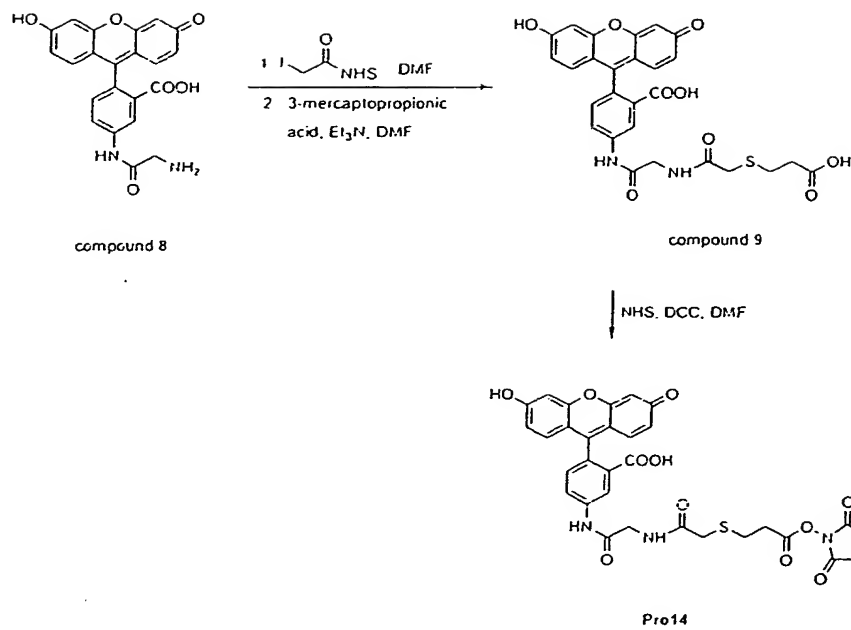
FIG. 10E

Synthesis of Pro15
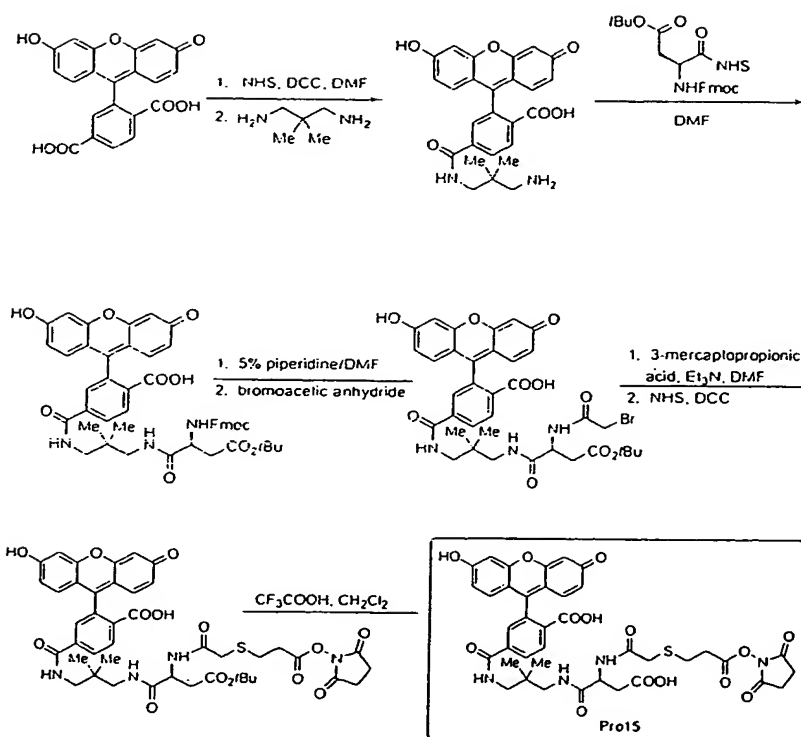
FIG. 10F

Synthesis of Pro20
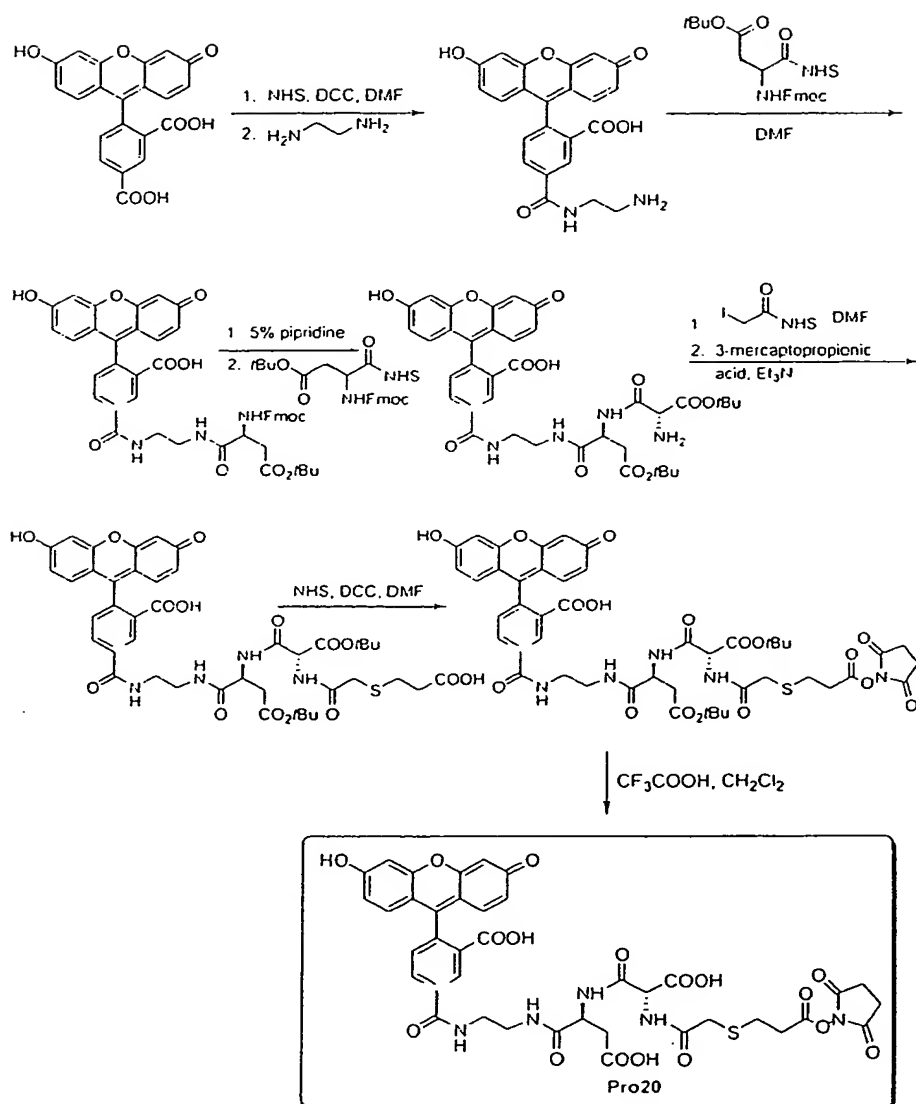
FIG. 10G

Synthesis of Pro22
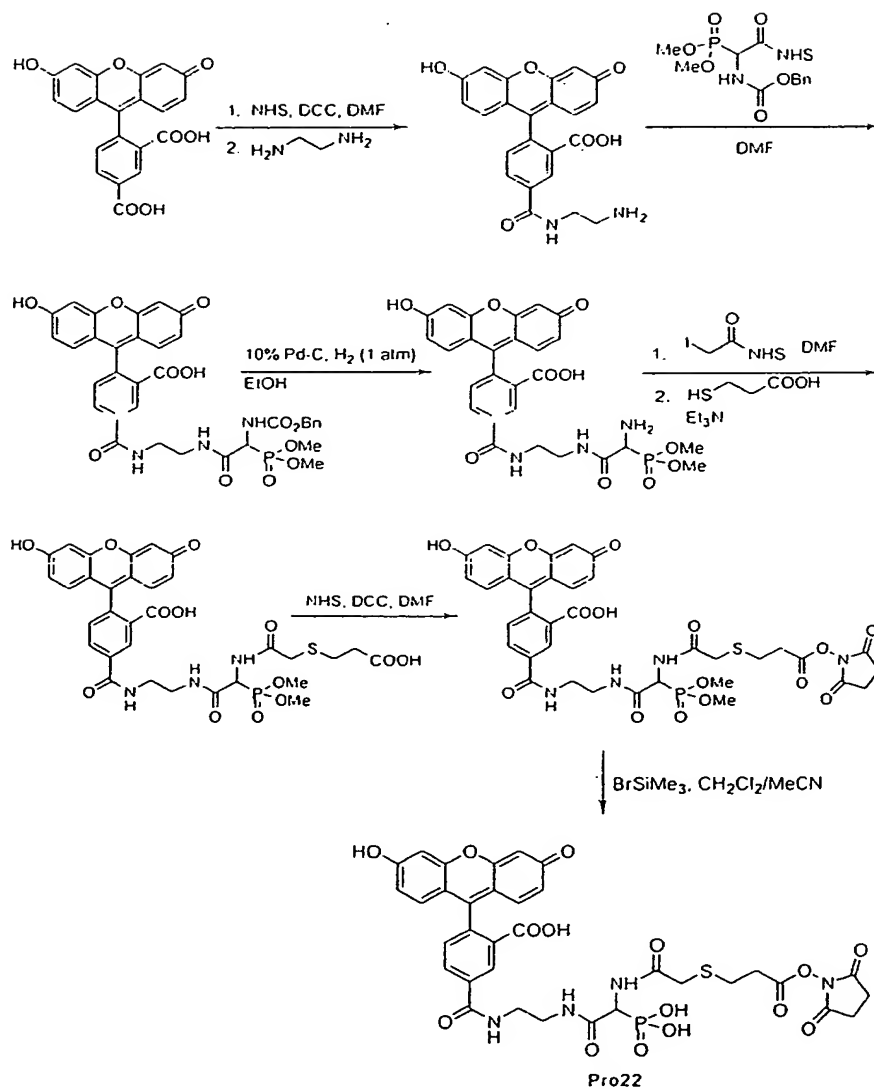
FIG. 10H

Synthesis of Pro28

FIG. 10i

Fig. 11A

Fig. 11B

Fig. 11C

Fig. 12A

Fig. 12B

Fig. 12C

Fig. 12D
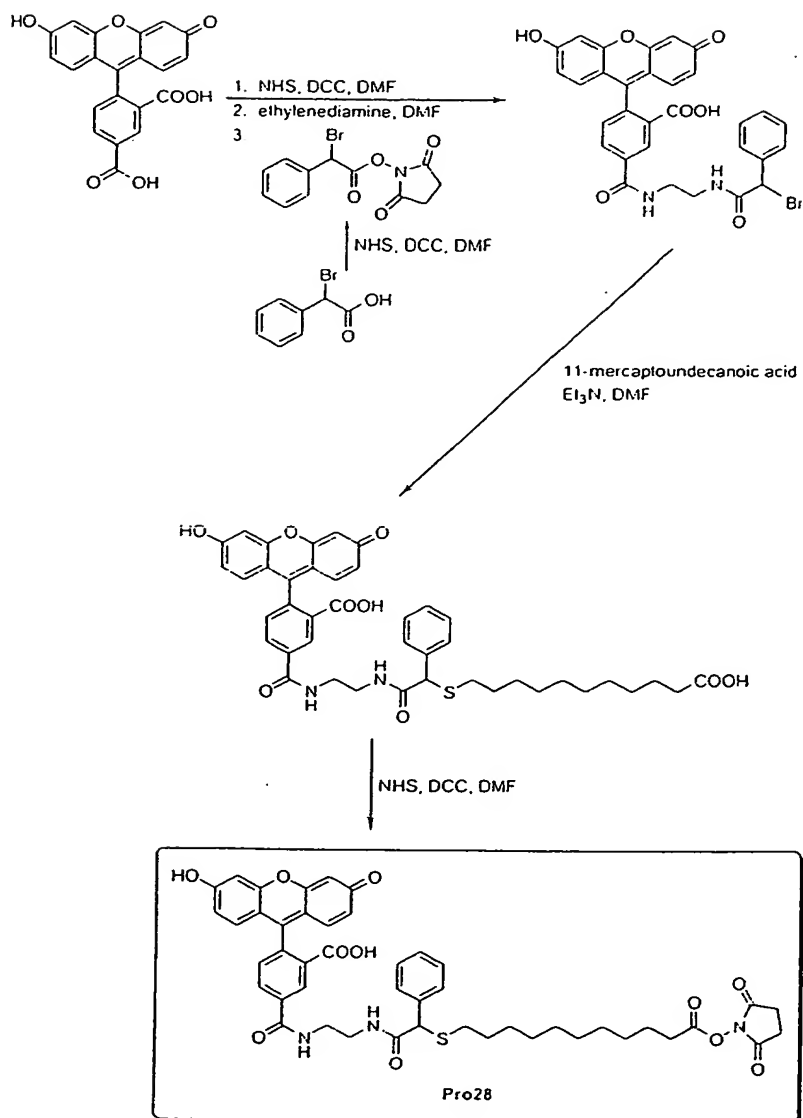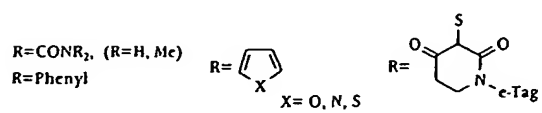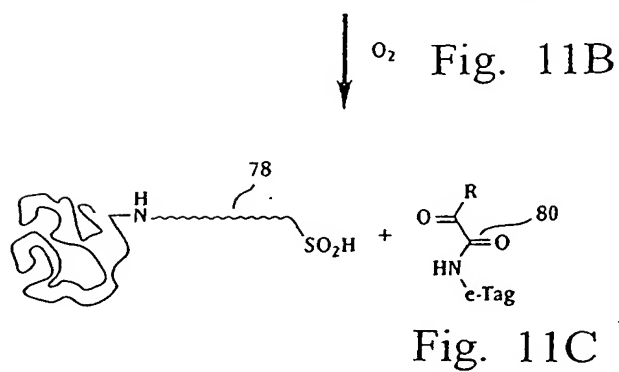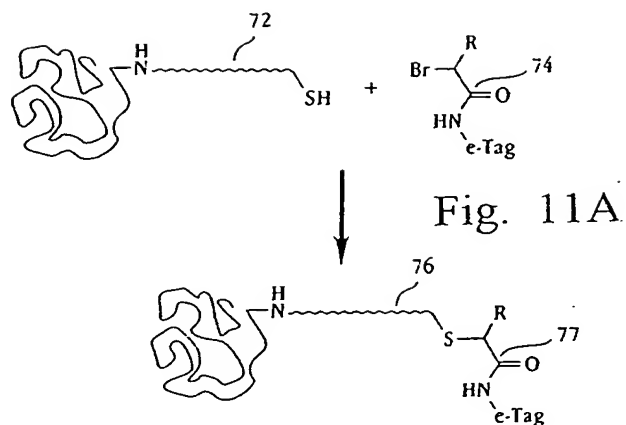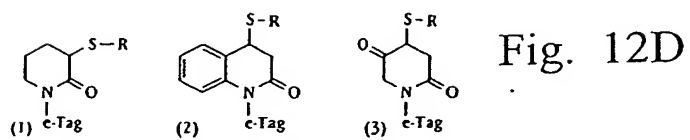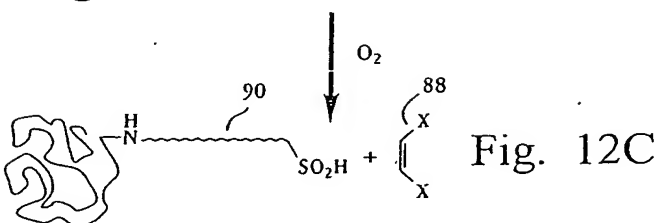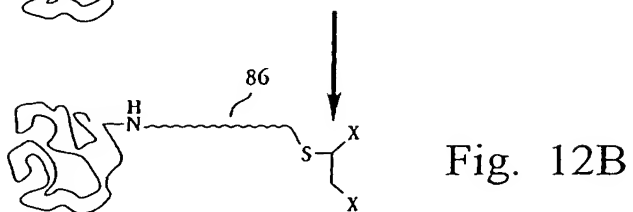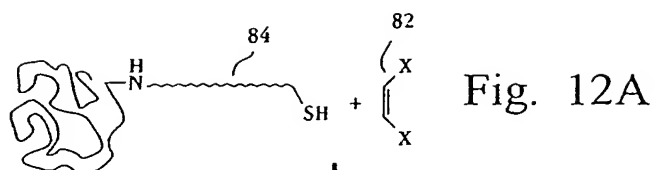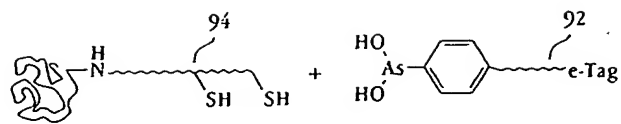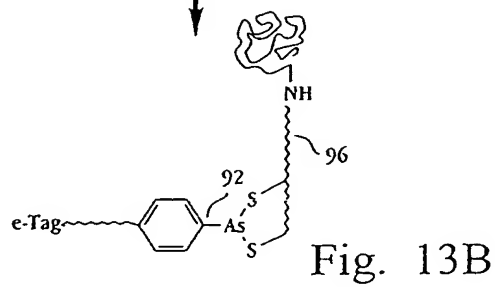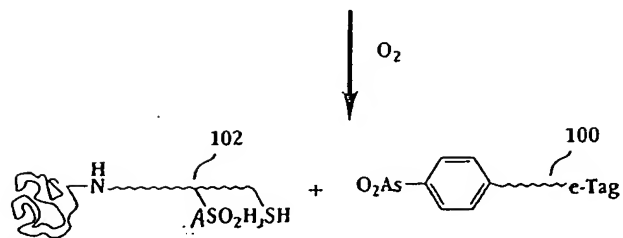
Fig. 13A

Fig. 13B

Fig. 13C
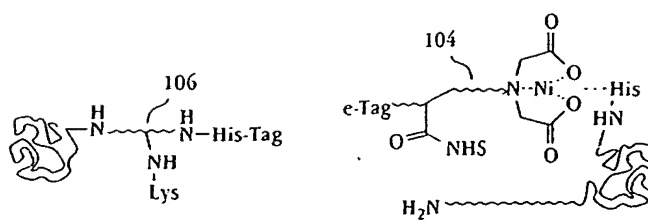
Fig. 14A
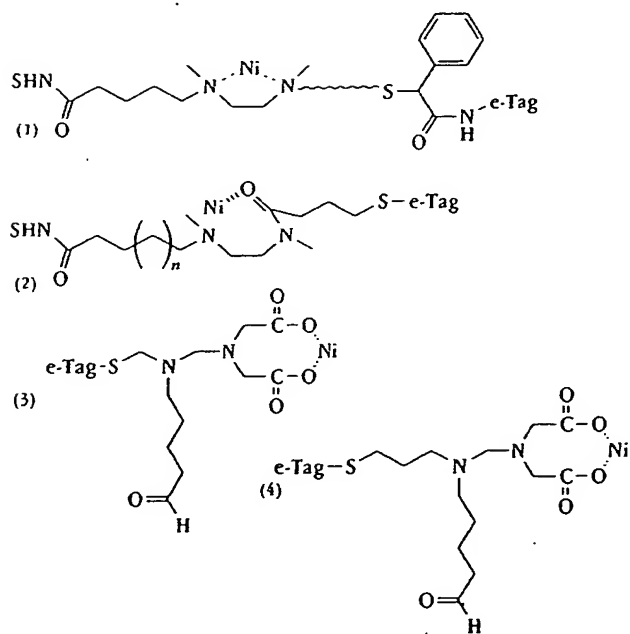
Fig. 14B
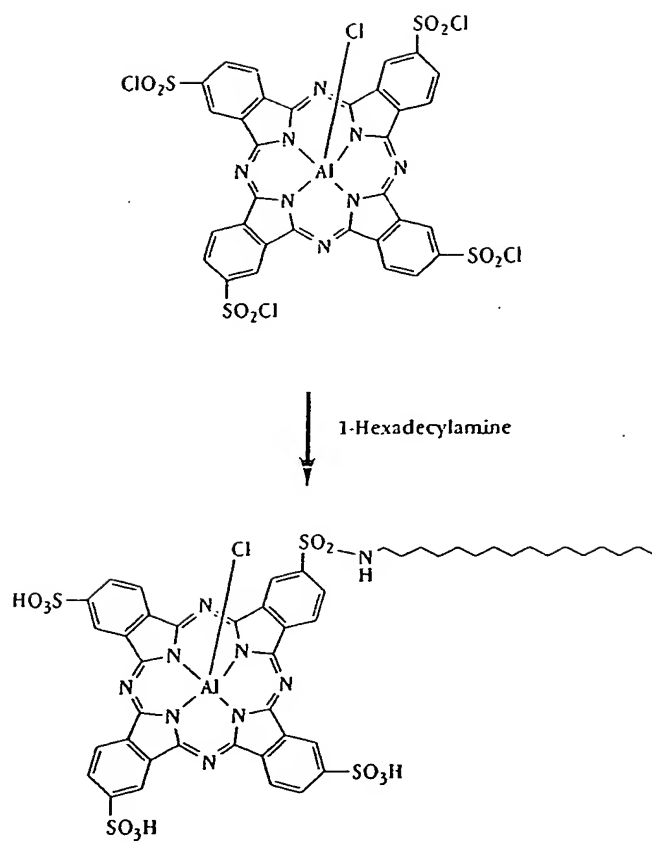
Fig. 15
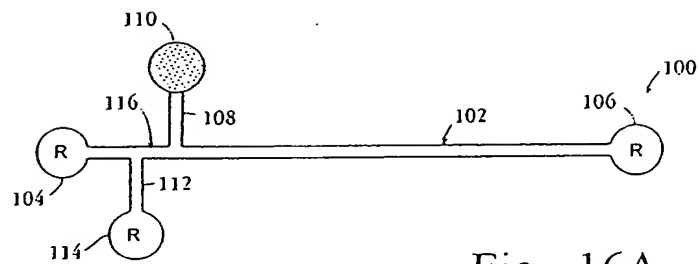
Fig. 16A
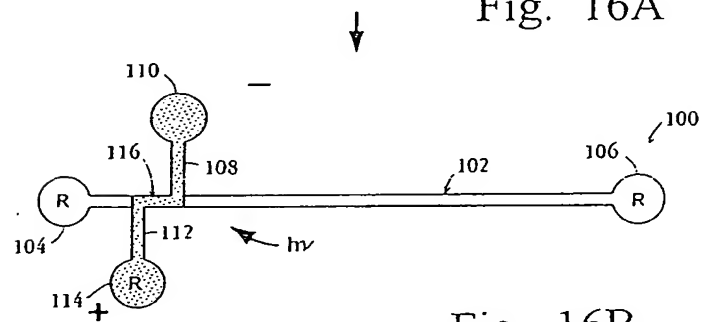
Fig. 16B
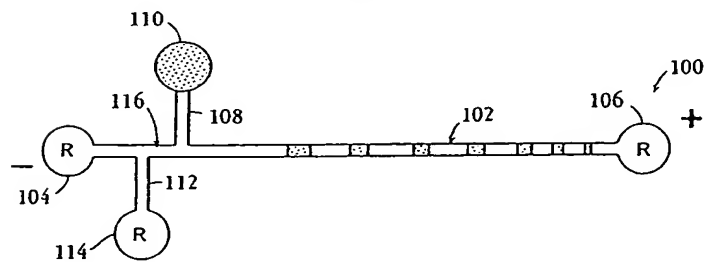
Fig. 16C
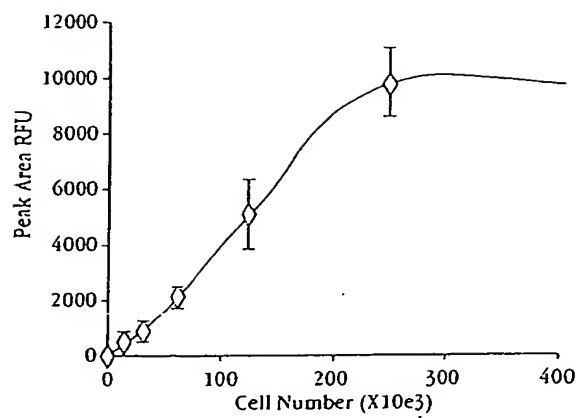
Fig. 17
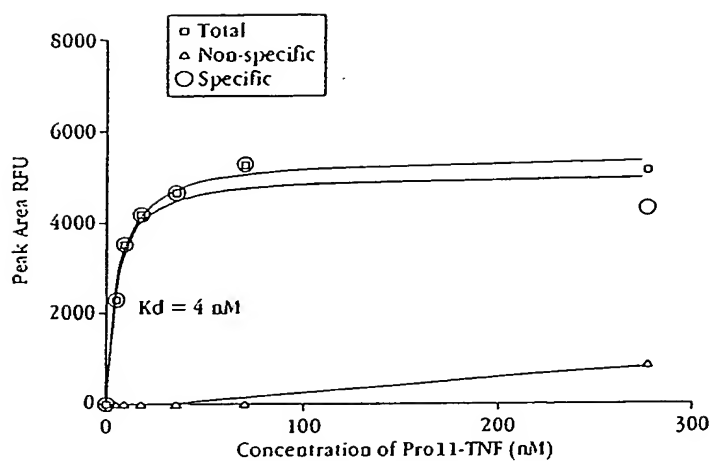
Fig. 18A
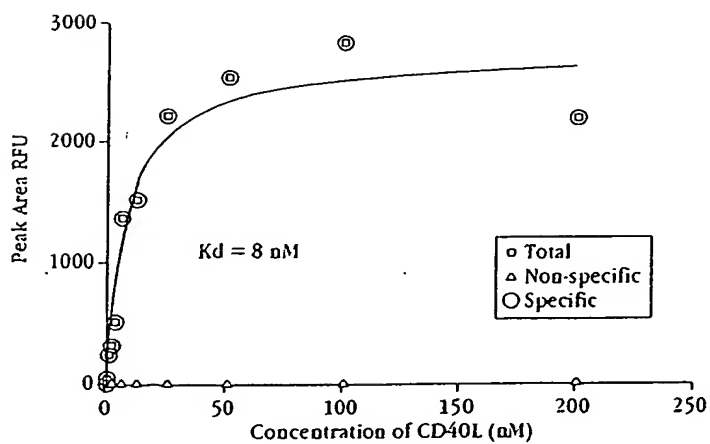
Fig. 18B
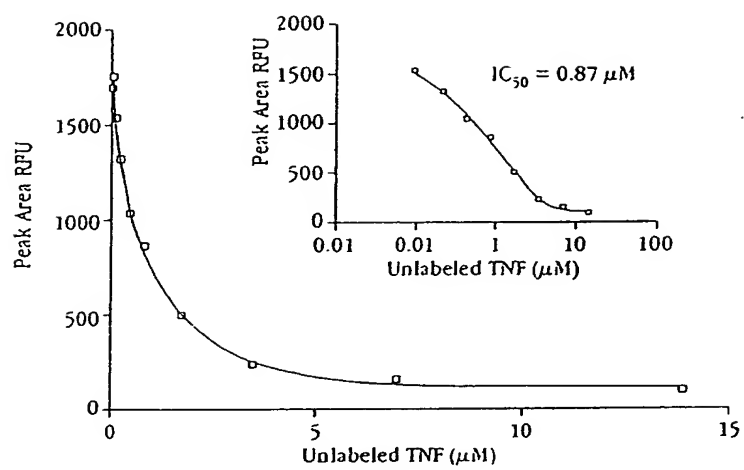
Fig. 19
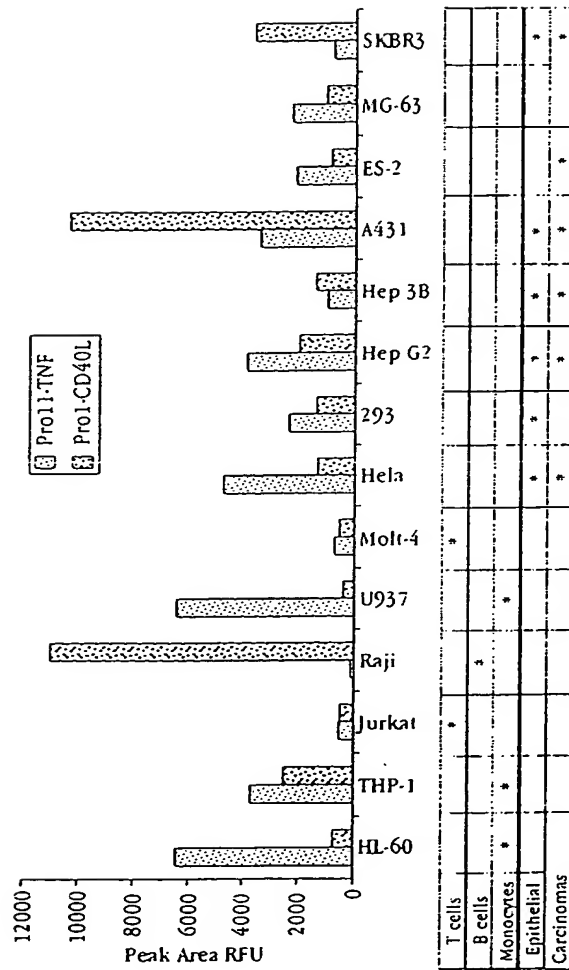
Fig. 20

Fig. 21A

Fig. 21B

Fig. 22A

Fig. 22B

Fig. 22C